

PRONOUN INTERPRETATION IN READING

by

EDWARD LIONEL WHITEHEAD

Plymouth Polytechnic

Thesis submitted to the Council for National Academic Awards
in partial fulfilment of the requirements for the degree
of Doctor of Philosophy

August, 1983

Acknowledgments

Steve Newstead, as Director of Studies, provided unstinting help and encouragement with all aspects of this research. Ian Dennis, as second supervisor, was particularly helpful with statistical and computing problems. I am grateful for all the trouble they have taken on my behalf. I am also grateful to Jonathan Evans, who generously read and commented on an earlier draft of this thesis.

The thesis itself was prepared by the author on the RUNOFF word processing system on Plymouth Polytechnic's PRIME computer. I would like to thank Steve Monk, of the Computer Centre staff, for the great care he took in printing the final version. I would also like to thank Suzanne Tolan for her diligent typing of the Appendices.

The experimental work presented in this thesis was conducted with the support of a Social Science Research Council studentship.

ABSTRACT

Pronoun Interpretation in Reading, by E. L. Whitehead

This research investigated the comprehension of anaphoric pronouns in reading. Chapter I explains the importance of this feature of text processing, and reviews the relevant psychological literature. One important feature of pronoun interpretation is the fact that antecedents vary in accessibility in response to aspects of text content and structure; antecedents which are readily accessible to pronominal reference are said to be foregrounded. The most common theoretical explanation of foregrounding is the claim that foregrounded antecedents are held in working memory. The review also shows reasons why good and poor readers might differ in sensitivity to foregrounding.

Experiments 1-3 investigated the effects on foregrounding of pronoun-antecedent distance and of topical continuity. It was found that distance had no effect on ease of pronoun assignment provided the intervening sentences were closely related to the antecedent. When the intervening sentences dealt with unrelated topics, pronoun interpretation became more difficult. The working memory explanation of foregrounding was investigated in Experiments 4-5, which found no evidence that foregrounding of an antecedent facilitated recognition memory for that antecedent. None of Experiments 1-5 found differences between the pronoun interpretation processes of good and poor readers.

There then follows a critical discussion of the working memory theory, which argues that foregrounding may also be explicable by a theory stressing antecedent retrieval processes. Later experiments attempted to discriminate between these theories. Experiments 6-7 found that antecedent foregrounding was sensitive to the content of the pronoun sentence itself. Experiments 8-9 found that foregrounding of an incorrect antecedent did not hamper pronoun interpretation any more than neutral backgrounding. All the results of Experiments 4-9 contradict the working memory theory.

It is concluded that the working memory theory of foregrounding is inadequate to explain the present results, and that an account of foregrounding based on antecedent retrieval processes is a plausible alternative.

CONTENTS

I.	Introduction and Literature Review	1
	1 Introduction	2
	2 Global Features of Comprehension	15
	3 Pronoun Interpretation	33
	4 Foregrounding	57
	5 Individual Differences in Reading Ability	84
II.	Distance and Foregrounding Effects on Pronoun Interpretation	97
	1 Experiment 1	98
	2 Experiment 2	121
	3 Experiment 3	158
	4 General Discussion	184
III.	Recognition Memory Tests of Foregrounding	189
	1 Experiment 4	190
	2 Experiment 5	211
	3 General Discussion	227
IV.	Alternative Explanations of Foregrounding	230
	1 Criticisms of the Working Memory Theory	231
	2 Retrieval Processes and Foregrounding	244
	3 Conclusion	261
V.	Pronoun Disambiguation and the Foregrounding Effect	265
	1 Experiment 6	266
	2 Experiment 7	295
	3 General Discussion	306
VI.	Processing Consequences of Incorrect Foregrounding	316
	1 Experiment 8	317
	2 Experiment 9	333
	3 General Discussion	349
VII.	Discussion	355

CHAPTER I

Introduction and Literature Review

- 1 Introduction
- 2 Global Features of Comprehension
- 3 Pronoun Interpretation
- 4 Foregrounding
- 5 Individual Differences in Reading Ability

Reading is a complex and varied cognitive skill which poses many interesting questions. Whether it is studied as a topic of interest in its own right, as an avenue of approach to a general understanding of human linguistic abilities, or as an applied problem with considerable practical relevance, the elaborate nature of the reading process offers a wealth of issues for investigation. This thesis is concerned with the specific issue of the interpretation of anaphoric pronouns; the present section, however, gives an overall survey of the nature of reading, and attempts to set the study of pronoun interpretation in its proper context.

One way of structuring the great complexity involved in reading is through a logical sequential ordering of what appear to be its necessary components. Thus, reading may be said to begin with visual pattern recognition, letter identification, and word recognition. Comprehension then requires a syntactic parsing mechanism to detect the meaningful word combinations, and a semantic interpretation mechanism to derive the meaning of the parsed units. Beyond the level of reading individual sentences, additional processes are required. These involve the identification of relationships between sentences, the drawing of inferences, and the integration of the various pieces of information into a unified representation of the text content. Such a set of processes, at progressively higher levels, could be said to constitute a characterisation of reading.

Research on reading has largely adopted the above approach, and by identifying distinct component processes has allowed each level to be

investigated separately in detail. Thus, there is a large literature on such topics as word recognition or syntactic parsing, as well as on text level processes such as pronoun interpretation. However, it is evident that no characterisation of reading could be complete without some explanation of the ways in which the different levels, isolated for research purposes, come together into a single overall process during comprehension. Three broad views of the way the different levels combine have been put forward. The first view stresses so-called bottom-up processing; by this account, the logical ordering of reading components is claimed to be the actual ordering during processing, with each level performing its appointed task and sending its output to be processed by the next highest level. This rigidly sequential view of reading has now been largely superseded. The second view of reading stresses the role of so-called top-down processes; these include the reader's use of knowledge and of expectations derived from higher levels of comprehension to direct the lower level processes. The top-down view has, however, been incorporated into the third approach to reading, which is currently the most widely accepted. On this view, exemplified by the explicit theory of Rumelhart (1977), the various levels of reading are interactively combined such that processes at any one level may affect processes at any other level. Thus, while information derived from the page may indicate that a certain word is being read, information based on semantic context or knowledge of the text topic may either confirm or contradict this identification. Information from each level is seen as being sent to a shared message centre (Rumelhart, 1977), where it becomes commonly available to all other levels. Again, the interactive view assumes that the different levels make use of shared central processing capacity; excessive processing load at one level may therefore reduce capacity available for processing at other levels.

The interactive view of reading is broadly accepted in the present research, and in later sections certain types of interaction will be referred to in discussions of different experiments. However, the reality of interactive processing does not invalidate the study of single reading components in isolation. While the dangers of possible distorting effects stemming from other aspects of reading must always be considered, it is nonetheless acceptable to focus specifically on one issue, and assume that the details of processing at other levels can be left on one side. Thus, in the present research the emphasis will be placed entirely on text level processes of identifying and using relationships between sentences, integrating information, and forming coherent text representations. On occasion, findings from studies of other aspects of reading will need to be considered; effects of syntactic structure on sentence processing, and of semantic context on word recognition, are two of these. However, the intention throughout is to elucidate the processes of pronoun interpretation, on the assumption that this subcomponent can be successfully included in a more comprehensive account of reading.

The importance of pronoun interpretation in reading is evident. While reading necessarily requires the understanding of each individual piece of information in a text, full comprehension is only attainable when these separate details are combined together to form a coherent representation of the text content. Pronominal reference is one way in which texts are unified into wholes. By using pronouns to refer to previously mentioned entities, a text producer creates direct links between sentences or between propositions within sentences; readers must detect these links between pronoun and antecedent in order to understand the text. In some sense, processes of pronoun interpretation may be common to both reading and listening; although the early,

low-level perceptual processes are clearly distinct, at a sufficiently higher level comprehension processes are presumably similar between the two modalities. Researchers on pronoun interpretation in reading have thus been able to make use of linguistic and psychological insights primarily based on speech, while at the same time results derived from reading research have a general applicability to human comprehension processes. The present research largely concentrates on such text-level processes, which may be common to reading and speech comprehension.

In psychology, the vast majority of research on pronoun interpretation has been conducted within the general processing framework exemplified by the given-new strategy of Clark and Haviland (1974, 1977). These authors accept the distinction between given and new information which has been identified by such linguists as Halliday (1970); given information in a sentence is that part of the content which is assumed to be already known to the reader or listener, while new information is the previously unknown content which the sentence conveys. Given information may be linguistically marked in several ways. One frequently mentioned marking is stress; given content is normally that which is spoken with only weak stress (Halliday, 1970). Other markings include the use of anaphoric expressions such as pronouns or definite noun phrases. According to Clark and Haviland, the importance of the given information in a sentence is that it provides a point of contact between the new information and the existing memory representation of the preceding discourse; in order for text content to be integrated into a coherent whole, each piece of new content must be added to memory at the most appropriate point. Clark and Haviland (1974, 1977) argue that there is an implicit cooperative contract between text producers and text understanders, such that the former deliberately structure successive sentences in such a way that their

marked given portion can be readily linked to what has gone before, while the latter make use of the assumed presence of such links when processing the text.

The given-new strategy makes two important assertions concerning pronoun assignment processes. Firstly, it is argued that readers attempt to identify the marked given information in sentences, and then make backward searches through their memory representation of the text in an attempt to find a matching antecedent. Thus, pronoun interpretation is seen as a process cued by encountering the pronoun itself. While it has been suggested (Garrod and Sanford, 1977) that readers may also engage in anticipatory checking of expressions as potential anaphors to certain prior concepts, the backward searching view of pronoun assignment has become generally accepted. The second important assertion of the given-new strategy is that pronoun assignment is a distinct process which precedes information integration; it is argued that readers first seek to identify the correct antecedent to given concepts, and only then attempt to integrate the new information at that point in the text representation. This separation between assignment and integration has also become widely accepted. However, the available descriptions of the given-new strategy are in fact quite vague regarding both the nature of the construction of a text representation and the nature of the antecedent search process; whether these two aspects of processing can be justifiably regarded as distinct stages may be open to question.

Evidence that readers attempt to form connections between given information and an earlier concept was presented by Haviland and Clark (1974), who found that sentences had longer reading times if the preceding context did not contain an explicit antecedent for their given

component. It was argued that in such cases readers make use of a variety of so-called bridging inferences to create a suitable antecedent representation. However, for present purposes the important questions concern the ways in which an antecedent is located in memory when one does in fact exist. In Clark and Haviland's model, it was assumed that the earlier mention of an antecedent would allow matches to be made, but no consideration was given to how such matches were determined by readers. The present thesis is specifically concerned with the influences which affect pronoun assignment.

There are many other aspects of text level processing in addition to anaphoric reference. For example, pronominal connection is neither a necessary nor a sufficient condition for textual coherence. Hobbs (1982) gives examples of referentially connected sentence sequences which do not form coherent texts, and proceeds to explore other sources of coherence than coreference; while Haberlandt and Bingham (1978) have demonstrated the role of verbs in forming sentence sequences into coherent narratives. Thus, pronoun interpretation must always be considered in the light of other forms of text integration. A second feature of text-level processing closely related to pronoun interpretation is the drawing of inferences; pronoun interpretation may be said to be a form of inference process. Readers do however form other types of inferential connection between sentences, such as relative spatial locations (Bransford, Barclay and Franks, 1972), or causal relationships between stated events (Kintsch, 1974). Ideally, pronoun interpretation would be accounted for in ways which showed the similarities between these different types of inference. A third feature of text processing which bears on anaphor interpretation is the use of pragmatic or world knowledge to guide comprehension processes. Schank and Abelson (1977) have argued that readers make continual use of

knowledge of stereotyped events, called scripts; this knowledge allows predictions to be made concerning likely future events in a text, or can be used to infer omitted features of a described situation. Script knowledge is relevant to anaphor interpretation by allowing references to be made to entities predicted by a script even if these have not been explicitly mentioned. A fourth aspect of text level processing, the construction of semantic macro-structures, is given extensive discussion in the following section of this chapter.

Pronoun interpretation is therefore embedded not merely within a set of comprehension processes occurring at different levels of reading, but also within a variety of processes occurring at its own level. This complexity poses problems for research on pronoun assignment; manipulations intended to affect pronoun interpretation in theoretically predicted ways must always be assessed for their potential confounding effects on other aspects of reading. In addition, theories of pronoun interpretation must be such as can be unified with what is known about other reading processes. However, while these constraints must be borne in mind, pronoun interpretation poses many specific problems which justify the attention it has received. The nature of pronoun interpretation is returned to below.

Working Memory

It is clear from the above brief outline of the overall characteristics of reading that one important feature of this process is the pooling of information. On the basis of the interactive view of reading, information from the different levels of processing must in some way be brought together to allow their mutual influences to occur;

furthermore, even within a single level of reading, different pieces of information must necessarily be combined. Thus, sentence parsing requires the appropriate combination of successive words, while text level processing by definition involves the linking together of separate sentences. Pronoun interpretation is one example of a text level process which requires the linking of two separate parts of the discourse, with the pronoun occurring somewhat later than the antecedent to which it must be connected. The generality of the problem of integrating separate parts of a text suggests that a general solution would be most appropriate, and it has been widely argued that the concept of working memory offers such a solution.

The concept of working memory has developed out of the earlier notion of a short-term memory store (Atkinson and Shiffrin, 1968), and shares some of its characteristics. Like short-term memory, working memory is assumed to contain a certain amount of limited storage capacity of short duration; in addition, however, working memory is seen as a central processing mechanism in which the various components of a cognitive task can be executed. This combination of storage and processing meets the requirements of information linkage in reading. For two pieces of information to be successfully integrated, it is assumed that they must be simultaneously present in working memory; items may enter working memory as a direct result of information extraction from the page, or may be reactivated from long-term memory. Once two items are present in working memory, comprehension processes may operate on them so as to combine them together or to draw on the two sources of information simultaneously. On these assumptions, comprehension strategies are seen as revolving around the appropriate control and selection of working memory contents in ways designed to facilitate understanding.

In reading, one of the most obvious uses of a working memory system is in sentence interpretation. While individual words can be understood in isolation, they cannot be combined into a unified sentence meaning unless the entire string is simultaneously available; sentence structure may in some cases not be completely apparent until relatively late in the sentence, and hence full understanding of the earlier words must be held in abeyance while the following words are read. In the sentence interpretation theory of Fodor, Bever and Garrett (1974), working memory is seen as being used for precisely this purpose. Early portions of what these authors refer to as a sentoid are held in working memory until the complete unit has been read; this unit can then be processed as a whole. The need to hold several partial sentoid representations in working memory is said to be responsible for the difficulty of comprehending self-embedded sentences; in such cases, working memory becomes overloaded before the full sentence is read. Jarvella (1971) produced evidence that verbatim representations of sentence content were available immediately after hearing a sentence but were lost once a further sentence boundary was crossed. Subjects listened to prose passages which were interrupted at certain points, when they were asked to make verbatim written recall of the immediately preceding content. Recall was found to be much greater and more accurate for the last sentence heard than for the preceding sentence; it was argued that syntactic boundaries were used in segmenting speech into manageable units which would not overload memory span. Presumably, similar considerations would also apply in reading.

While the role of some form of working memory system in reading is frequently asserted, only rarely is any attempt made to refine or specify this concept. Reading researchers generally point out that some such mechanism would serve several useful purposes, and may explain

various findings in the study of comprehension on that basis, but do not investigate the properties of that system in its own right. Probably the most detailed attempt to understand the nature of working memory has come within the framework originally presented by Baddeley and Hitch (1974), and specifically extended to problems of reading by Baddeley (1979).

In the Baddeley theory of working memory, this system is not an undifferentiated mechanism but is divided into several distinct subsystems. One of these, termed the central executive, is seen as possessing limited storage capacity of its own but as also being the controlling processor; as such, it appears in many ways to be virtually a working memory within working memory, and since this is the original concept which was under investigation it adds little to what is known. However, the Baddeley model has also suggested that working memory includes separate slave subsystems, under the control of the central executive, which provide additional storage capacity of a distinct type. The slave system which has received most attention is termed the articulatory loop. This is conceived as a limited capacity and phonologically coded store which is responsible for such phenomena as phonemic confusions in short-term memory tasks. Evidence for the role of the articulatory loop in reading comes from results such as those of Baddeley, Eldridge and Lewis (1981), who required subjects to judge whether presented sentences were semantically acceptable or anomalous; accuracy of identification of anomalous sentences was significantly poorer when subjects were also required to count aloud from one to six while performing the task, which presumably interfered with use of the articulatory loop. However, since speed of responding was unaffected, and since accuracy of identification of acceptable sentences was also unchanged by counting, it was concluded that the loop acts at most as a

supplementary source of storage which would be of most use only in difficult reading tasks which required retention of order information. Levy (1978) similarly concluded that phonological recoding in working memory was not essential for semantic processing.

While the role of the articulatory loop in skilled, fluent reading is likely to be small, Baddeley (1979) argued that it may be of greater relevance in learning to read. One line of evidence put forward in support of this view was the fact that effects of phonemic similarity only begin to appear at about the age when children are learning to read; it was argued that the loop might be used in blending individual speech sounds into syllables, and by relieving the load on the central executive would thus facilitate processes of word encoding. The fact that poor readers show weaker effects of phonemic similarity than good readers was also said to indicate that they made less use of the loop, and so were burdening the central processor with tasks which could be more efficiently delegated to the slave subsystem.

It is evident that the main uses of such a system as the articulatory loop would be at the word and sentence level, rather than at the text level. While this system may assist in word encoding, or may possibly be the system used to maintain early parts of a sentence until later parts have been read, it is not obvious that phonologically coded storage of this type would be of use in tasks such as pronoun assignment or drawing of inferences. These tasks primarily require use of semantically coded information. In addition, if the loop is occupied with maintaining information for use in lower level processes, its availability to higher level processes must be reduced. Possibly the most interesting suggestion concerning the loop is that its use may free central processing capacity for other tasks; thus, if text level

processes are primarily carried out by the central executive, use of the loop for lower level processes may be indirectly facilitative. However, this leaves the burden of the explanation of text level processes on the central executive, which is admitted to be poorly understood at present (e.g. Baddeley, 1981). As noted above, the central executive is in fact attributed with many of those properties which were originally attributed to the undifferentiated working memory system as a whole; separating out slave subsystems such as the articulatory loop or visuo-spatial scratch pad (Baddeley and Lieberman, 1980) has left that central executive unexplored.

In essence, the concept of working memory is used by reading researchers as an explanation for patterns of item availability during processing. While most theories assume a structural separation between working memory and long-term memory, it has also been suggested (Anderson and Bower, 1973) that working memory is simply a currently active partition of long-term memory. As will be described in a later section, theories of pronoun interpretation in particular make heavy use of the concept of working memory, by attributing ease of interpretation to the ready availability of antecedents in that storage system. The general assumption is that working memory contents are under strategic control in ways designed to maximise comprehension efficiency. In the absence of a meaningful model of working memory, including details of its capacity, the type of coding used, and possible structuring of the information represented there, this assumption would be difficult to test empirically; the definition of working memory as a storage system in which items are readily accessible to processing, and the assertion that readily accessible items must therefore be in working memory, in fact says little about the relevant comprehension processes. However, certain general predictions based on the notion of working memory

presence may prove to be testable; for example, if working memory is indeed a shared central mechanism, items present there would necessarily be available to all processing tasks rather than to one specific task. Predictions such as this do not rest upon details of the nature of working memory, but on broadly accepted assumptions concerning its role in processing.

In later sections of this thesis, the appropriacy of working memory presence as an explanation for certain features of reading will be put to such general tests as this. However, in common with the majority of research on pronoun interpretation, the studies presented here will not attempt a detailed characterisation of the nature of working memory.

A widespread observation in the study of reading is that there are at least two distinct levels of processing. At one level, readers must identify words, parse sentences, and identify connections between sentences; the processes of pronoun interpretation are normally conceived as part of this lower-level comprehension. However, at a higher level readers must process the content of a text in terms of its overall meaning and structure, identifying the relevance of and relationships between text units which may be considerably larger than sentences or pairs of sentences. These more global aspects of text processing are of interest in their own right, but it will be argued in a later section that they also have relevance for the relatively lower-level processes of pronoun interpretation. This review will deal with two theoretical approaches to higher-level features of comprehension; the reduction of text content into semantic macro-structures, and the abstract constituent structure of texts.

The theory of semantic macro-structures was originally put forward by van Dijk (1977a, 1977b) in order to account for such skills as the ability to identify discourse topics or to write summaries of texts. It was argued that these abilities cannot be explained solely by considering the sequence of sentences which a text contains, but required an understanding of the ways in which sets of sentences might be related together and reduced in content. Such reduction of information is an essential feature of comprehension of longer texts, which evidently are not retained verbatim. However, van Dijk's initial approach to the question of information reduction was a purely formal one; although intended to underpin processing accounts of known

features of comprehension, macro-structure theory itself did not contain processing claims. That theory will now be described, followed by a review of later attempts to embody that theory in a psychological model.

A characteristic feature of macro-structure theory, and of related processing models, is the use made of propositional representations of text content. The surface sentences of a text may each express several propositions. A proposition is a unit of meaning composed of a predicate or relational term together with one or more arguments; predicates are such concepts as may be expressed in surface verbs or adjectives, while arguments are normally noun concepts or other propositions. The sequence of propositions representing a text is termed the text base; the propositions of the text base are ordered according to the order of occurrence of the surface items which they represent. A crucial requirement of the theory is that any text base which is to be used in macro-structure formation must be coherent. Coherence consists in essence of argument repetition between propositions; that is, two propositions are said to be coherent if they share at least one argument in common. It is not necessary that each proposition in the text base include an argument which is common to all the others, since for example three propositions are coherent if the first and second have one argument in common while the second and third have a different argument in common. Anaphoric reference is a form of argument repetition which is reflected in surface sentences. However, the surface text may not in fact make all such argument connections explicit; definite noun phrases may be used, for example, without any prior mention of the specific concept being referred to. This difficulty is overcome by drawing a distinction between the explicit text base and the implicit text base. The implicit text base is the

sequence of propositions which is directly expressed in the surface sentences of a text; the explicit text base is a theoretical construct which has been made fully coherent by such means as supplying omitted antecedents.

Macro-structure formation therefore operates on an explicit text base of so-called micro-propositions; according to van Dijk, the macro-structure of a text is entailed by these micro-propositions. In general, macro-structures are derived by replacing sets of micro-propositions with a single macro-proposition which expresses their most important content but omits irrelevant detail. Like the sequence of micro-propositions, the sequence of macro-propositions must itself be coherent. There may in fact be several levels of macro-propositions, since the processes of macro-structure formation are recursive; a sequence of macro-propositions derived directly from the explicit text base can itself be reduced into a sequence of higher-level macro-propositions. Each successively higher macro-proposition subsumes larger sections of the original text content. The derivation of macro-propositions occurs by the application of macro-rules; each such rule operates in a manner designed to reduce the content of the text or to express the most important and relevant pieces of information. Four such rules are proposed by van Dijk (1977a), and although later formulations have made slight adjustments the overall manner of their operation has remained unchanged.

The first macro-rule is termed Generalisation. This rule takes sets of concepts and replaces them with a reference to a shared superordinate concept. The rule applies to both predicate and argument concepts. As examples, van Dijk (1977a) states that references to 'a dog' and 'a cat' may be generalised into a macro-proposition referring

to 'pets', and that statements concerned with 'flying to Paris' and 'taking a train to Paris' may be generalised into a macro-proposition dealing with 'going to Paris'. Generalisation may be applied simultaneously to the predicates and arguments of a set of propositions; van Dijk's example shows how statements describing the different activities of a mother and father and their children may be generalised into a single statement that the whole family was working. Generalisation therefore abstracts from unnecessary detail. The second macro-rule, however, termed Deletion, removes full propositions from the text without reducing their content into a higher-level macro-proposition. Such details are removed due to their lack of relevance to the main point of the text; thus, from a description of a child breaking a window with a blue ball, the proposition stating that the ball was blue could be deleted. The converse of the macro-rule Deletion is termed Integration; this rule takes micro-propositions from the text base and adds them to the macro-structure unchanged. Such integration reflects the fact that texts may make explicit statements that are relevant at the macro-level. Thus, a statement that someone took a train to Paris may be followed by descriptions of the predictable details of buying a ticket or boarding the train; these details may be subjected to Deletion, while the proposition expressing the main point of those details is carried directly into the next level of macro-structure. Finally, the macro-rule of Construction deals with cases where predictable details of an activity are described but no explicit statement of their global import is contained in the text. As an example, it is stated that descriptions of someone going to a railway station and buying a ticket need not be accompanied in the text base by a proposition stating that the person took a train; yet such a macro-proposition would be constructed on the basis of the explicit details.

This is only a brief and informal description of the macro-rules, yet their overall nature and purpose are evident. The rules operate on an explicit and detailed text base, and reduce and summarise its content to the main topical points. Van Dijk (1977a, 1977b) attempts to present the rules in a formal notation, with explicitly stated conditions and constraints upon them; yet it is apparent from this attempt at rigour that many aspects of the operation of the rules are left unexplained. For example, the choice of the superordinate term to replace sets of micro-propositions during Generalisation must be constrained by such things as the purpose of the text or the nature of the intended audience; 'dog and cat' may be justifiably generalised to 'mammals' in a biology lecture. Again, the identification of predictable components of global activities, which is crucial to Integration and Construction, requires the use of much complex world knowledge; van Dijk (1977a) stresses the use of script-like representations to guide this process. However, the most serious difficulty with the approach is the frequent use of the concepts of 'irrelevance' and 'global importance'. Details can be deleted or generalised if they are irrelevant, while macro-propositions can be integrated or constructed if they are globally significant; thus, the derivation of macro-structures to express global content is dependent on the use of prior assessments of such relevance. No formal manner of identifying relevance is provided, and the derivation of macro-structures is usually accomplished in practice through the researcher's intuition. Van Dijk recognises the need for some definition of irrelevance, and states as a constraint on Deletion and Generalisation that no proposition may be omitted from the macro-structure if it is required as an interpretation condition for some later micro-proposition; this restriction is admissible in the theory of macro-structures, since it is assumed that the entire text base is simultaneously available to the macro-rules. An overall

constraint of this type is evidently necessary, but to be systematically applied it would require a full statement of what may or may not be an interpretation condition for a later proposition; clearly, one interpretation condition for an anaphoric expression must be the prior mention of its antecedent, but this is presumably not the only form of condition which exists.

The many omissions in the formal theory of macro-structures are not damaging to the merit of this overall approach. This is partly because the reduction and organisation of information play an undoubted role in text comprehension and memory, and therefore the general features of the model have relevance to any processing account of reading; empirical evidence in support of this model will be reviewed shortly. The organisation of text into subunits, and the labelling of those subunits with statements of their overall content, is an important theoretical contribution of the macro-structure approach. In addition, the macro-structure approach is valuable because the attempted rigour, however inadequate current versions may be, does at least enable the points of vagueness to be identified, while those aspects of the theory which have been relatively fully specified can be put to empirical test. Other approaches to text comprehension and representation, such as the discourse model theory supported by Webber (1980) and Johnson-Laird (1981), are no less incomplete yet do not as yet possess these heuristic advantages.

The formal theory of macro-structures, although not itself a processing model, forms part of the psychological theory of reading presented by Kintsch and van Dijk (1978; also Kintsch and Vipond, 1980). This model has several interesting features which give it relevance to the review of pronoun interpretation which is presented in

the next section, and it will therefore be described in full rather than solely with respect to the processing of macro-structures. Crucially, this model emphasises the construction of a referentially coherent text representation, and makes specific assertions as to how this is achieved; the model also makes extensive use of the concept of working memory.

One problem with using the theory of macro-structures as part of a processing model of reading is its assumption that the entire text base is simultaneously available; this is evidently not the case for readers processing a text sequentially. In recognition of this, the Kintsch model assumes that readers process a text in a succession of input cycles, composed of several propositions each. The number of propositions taken in during each cycle varies, and is responsive to sentence and clause boundaries; each input cycle normally processes one or more syntactic units, depending on the number of propositions which those units express. During the first cycle of reading a text, the input propositions are organised into a coherence graph structured by the argument repetitions between propositions. Coherence graphs are represented as hierarchical trees, with each node representing a proposition and links between nodes showing which propositions share an argument. The topmost propositions of such a graph are said to be presuppositions of the propositions directly subordinate to them, since they introduce discourse referents which the lower propositions repeat; a coherence graph may have several levels of propositions, with argument repetition connections between each level.

Once a coherence graph has been created out of the first input cycle of propositions, readers may proceed to the next input cycle. However, if each input cycle were processed separately and used to

create a distinct coherence graph, it is clear that the text representation as a whole would not be coherent. To ensure that successive sets of propositions are linked together, the model assumes that certain selected propositions are carried forward from one cycle to be used in processing the next. These selected propositions are carried forward in a short-term buffer of limited capacity. During processing of the following cycle, an attempt is made to connect the input propositions with those held in the buffer store; if argument repetitions allow such connections to be made, the resulting graph will not only be internally coherent but will also be linked to the propositions from earlier cycles which were not carried forward. Difficulties arise, however, if the input propositions cannot be connected to the propositions in the buffer. In such cases, reinstatement searches must be made of the long-term memory representation, in an attempt to locate an earlier proposition which was not originally selected for the buffer but to which the new input can be connected. If such a proposition is found, it will be reinstated to working memory and construction of a coherence graph will proceed once more. Evidently, such reinstatement searches would be expected to cause processing difficulty compared with direct connection of input propositions to propositions already in the buffer. In certain circumstances, the model accepts that it may prove impossible to interconnect all the propositions of the text, and hence the eventual text representation will consist of two or more separate coherence graphs; these discontinuities may be resolved by inferences of explicitly unstated argument connections.

The ease with which a coherence graph can be constructed is thus dependent on the presence of interconnections between the propositions selected for buffer storage and the input propositions of the following

cycle. The likelihood that such connections will exist depends on the nature of the selection strategy which determines the propositions to be carried forward from one cycle to the next. According to the Kintsch model, propositions are selected for the buffer on the basis of importance and recency. Important propositions are those which occur high in the existing coherence graph and have many connections to other propositions; a proposition with many connections is more likely to be relevant to the next cycle than a proposition with few connections. Recency, as the name implies, favours the propositions which have most recently occurred in the text. By combining these two criteria, the selection strategy will choose propositions for the buffer that are most likely to be connected to the next set of input propositions. As will be discussed in the next section, the efficient selection of working memory contents has played a central role in theories of pronoun interpretation. However, in the Kintsch model the two criteria given above do not in fact define a unique selection strategy; various ways of specifying the criteria and their manner of combination would lead to different predictions concerning which propositions were selected for the buffer. The decision between the different possible strategies is seen as an empirical one, but most research carried out within the framework of this model has assumed the validity of the so-called leading edge strategy. This in essence selects the set of propositions along the edge of the constructed coherence graph extending from the topmost proposition to the most recent lowest-level proposition.

The model therefore sees texts as being processed in cycles of several propositions, with a selected subset of propositions being carried over from one cycle to the next. In addition to the nature of the buffer selection strategy, two other crucial parameters of the model are the size of the buffer itself and the number of propositions taken

in during each cycle. Both of these affect the ease with which a coherence graph can be built up. If the buffer size is small, few propositions can be carried forward; while if the input cycle is small, connections to the buffer are less likely to be present. In this way, the values given to these parameters affect the number of reinstatement searches which processing a particular text will require. The size of the buffer is a consideration which is particularly relevant to the present research, since it has been suggested by Kintsch and his associates that less-skilled readers may suffer from small buffer capacity; this theory of reading ability is discussed in a later section. The value set to the two parameters not only affects ease of comprehension, but is also likely to affect patterns of recall. The model predicts that recall probability for each text base proposition is dependent on the number of times it has been selected for buffer retention and featured in a further processing cycle; the more cycles in which a proposition has been processed, the greater the strength of its coding in memory and the greater the likelihood that it will subsequently be recalled.

The above description of the Kintsch model, based on the presentations in Kintsch and van Dijk (1978) and Kintsch and Vipond (1980), has been concerned solely with the micro-processes of comprehension involved in forming text base propositions into a coherent representation. Unfortunately, the operation of the macro-processes of comprehension has been given less attention in the existing literature. According to Kintsch and van Dijk (1978), macro-processes occur simultaneously with micro-processes, with the buffer store being used to retain relevant macro-propositions and other information required to establish global coherence; however, the cyclic operation of macro-structure processes, and the difficulties caused by being unable

to inspect an entire text base simultaneously, have not been extensively discussed. The most useful observations concerning the actual operation of macro-processes have been made by van Dijk (1977a), who suggests a number of strategies which readers may use to create macro-structure representations. One feature of these strategies is that they are probabilistic and predictive; readers are said to construct hypotheses concerning plausible macro-structures, before all the necessary text has in fact been read. Such hypotheses may later be disconfirmed, but this approach accepts the reality of sequential text processing.

One suggested strategy is based on inferences from explicitly mentioned components of standard activities. Thus, mention of someone going to a railway station may lead to a hypothesis that he will take a train; this would form the assumed topic of the section of text, and be the basis for macro-structure formation. A second strategy is the use of explicit macro-structural discourse cues. These include titles, abstracts, introductory statements of the purpose of a text, and topic marker expressions such as 'the crucial point is...'. Other strategies make use of general knowledge, such as by predicting the likely main points of an article on the oil crisis. Kintsch (1977) suggests that texts may also be segmented into macro-structural units on the basis of paragraph divisions; he also argues for the importance of discontinuities in the content of the text, such as changes of time, of location, or of the characters currently involved. All such cues will be useful in determining macro-structure boundaries as a text is read.

Clearly, the above suggestions are inadequate as a model of macro-structure formation during reading, but the general insights which they offer can be readily refined and put to empirical test. At

present, however, the primary evidence for the psychological reality of macro-structures comes from studies of text memory. Kintsch and van Dijk (1978) tested their overall reading model by presenting subjects with a text describing a psychological investigation, and obtaining recall protocols either immediately or after a delay of one or three months; subjects also gave summaries, after writing their full recalls. The experimental text was analysed on the basis of the model, using specified assumptions concerning the size of the buffer and of each input cycle, the nature of the selection strategy, and the recall consequences of each process that was applied to each proposition. As noted above, recall probability of a micro-proposition was assumed to increase each time it entered in a new processing cycle; recall probability of a macro-proposition was seen as a function of the number of successive levels of the macro-structure into which it was integrated. This macro-structure was derived by intuitive application of the macro-rules to the text base. Although the model specified the ways in which recall probabilities for each type of proposition would be affected by the processes they went through, it did not specify the size of these effects; the effect on recall probability was to be estimated by statistical comparisons between the formal model and the experimental data.

It was found that the overall fit between the predictions based on the model and the obtained recall probabilities was reasonably close. In addition, the pattern of recall probabilities showed clear differences between micro- and macro-propositions. The recall probability of micro-propositions fell considerably when recall was delayed, but recall of macro-propositions was relatively unaffected by delay. This outcome is consistent with the view that macro-propositions are used to organise text memory, and represent the main content of the

text, while micro-propositions are recalled mainly by using the macro-structure as a retrieval cue. Also supporting the view that the formally derived macro-structure represents the main content of a text was the finding that subjects' summaries were composed mainly of macro-propositions. Further evidence for the reality of macro-processing was found when the model was tested against recall protocols from a condition in which subjects had read only the first paragraph of the analysed text; reading such a fragment in isolation should prevent the operation of macro-processes. On this occasion, the processing model made only a poor fit to the data, while recall probabilities did not differ between micro- and macro-propositions. Thus, preventing subjects from using macro-processes reduced the validity of the model.

Van Dijk and Kintsch (1977) briefly refer to similar results obtained using narrative texts rather than a psychological report. Here again, recall protocols consisted solely of macro-propositions if recall was delayed, but included micro-propositions if recall was immediate; summaries again appeared to be based directly on the macro-structure. Miller and Kintsch (1980) used a procedure similar to that of Kintsch and van Dijk (1978), with short passages and immediate recall only, in order to minimise subjects' macro-processing; here too the overall closeness of the fit between the model and the obtained data was reduced. Thus, the main results of Kintsch and van Dijk (1978) have been replicated elsewhere, on different types of passage. Kintsch and Yarbrough (1982) showed the importance of explicit cues to text structure by presenting subjects with passages in which such cues either were or were not available; memory was tested either with a cloze procedure, which was intended to reflect successful execution of micro-processes, or with questions asking for the topic and main points

of the passage, which were expected to reflect subjects' success at forming macro-structures. It was found that cloze performance was unaffected by the presence of cues to structure, but that question performance was significantly worse when this structure had been made difficult to detect. This result not only confirms the separability of micro-processes and macro-processes, but also showed the value of certain cues to macro-structure formation.

Vipond (1980) made a detailed study of the separability of micro- and macro- processes. This study is interesting because it presents a cyclical model of macro-processing directly analogous to that already described for the processes dealing with more local coherence. Macro-processing is seen as occurring in input cycles of paragraph size, with macro-propositions organised into a hierarchical network based on argument repetition, and with selected propositions carried over to the next cycle in a distinct macro-buffer. This extension of the cyclic model nonetheless omits any full treatment of the derivation of macro-propositions from the text base; this was carried out by the experimenter on intuitive grounds. Using the cyclic model, texts were analysed and predictions made concerning their ease of processing at the two levels; as noted above, one influence on ease of processing is said to be the number of reinstatements required to establish coherence, although Vipond used other predictor variables as well. The model's predictions were then compared against subjects' reading times and recall performance on the analysed passages. It was found by factor analysis that the predictor variables at the micro- and macro-levels fell into distinct clusters. Further experiments showed that when micro-processing difficulty and macro-processing difficulty were orthogonally varied, no interaction between their effects was found; that is, greater text difficulty at one level did not prevent subjects'

taking advantage of ease of processing at the other level. This result was taken to show that the two types of processing are independent, and do not make use of shared cognitive resources, although full explanation of ease of comprehension required both levels to be taken into consideration.

There is thus considerable research evidence demonstrating the importance of macro-structural processing during comprehension. While this research has not yet shown how readers in fact create macro-structure representations during comprehension, the effects of variables likely to influence the ease of such construction argues for the reality of that form of processing; and the correspondence between macro-structural text analyses and subjects' recall performance again shows that, however this level of representation is derived in reading, the outcome of those processes has measurable consequences.

The role of macro-structures in comprehension can therefore be summarised as follows. The sequence of sentences making up a text is inadequate as a representation of the meaning of that text; for full appreciation of a text's content, it is necessary to combine sets of sentences together into higher level units expressing globally relevant information. These units are labelled with macro-propositions expressing that content. Such macro-propositions initially contain all the information needed in the interpretation of later micro-propositions, but omit irrelevant detail; recursive application of macro-processes, however, can create still more general levels of the macro-structure. Information which is deleted from the macro-structure is not seen as being deleted from memory, since micro-propositions at low levels may also be retained; but memory is organised in terms of a text's macro-structure, and macro-propositions are less sensitive to

forgetting with delay. In attempted recall, the macro-structure is seen as functioning as a retrieval cue assisting readers in their recall of lower-level propositions. Such propositions may be recalled directly from memory traces subordinate to macro-propositions, or may be inferred by inverse application of the macro-rules.

One additional feature of macro-structure formation must also be briefly mentioned. Kintsch and his co-workers argue that the macro-structure of a particular text is always constructed within an overall controlling schema; this schema may be derived from such things as the reader's task or goals, which will influence decisions as to which pieces of information are relevant and which may be deleted from the macro-structure, or alternatively the schema may be derived from the abstract structural organisation of the text. It has been widely observed that texts may be analysed into constituent categories, and relations between categories, that are defined not by specific surface content but by restrictions on the type of content which they may contain. These predictable variations in type of content between different parts of a text have important consequences for macro-structure formation, since they can be used to define the types of information reduction that are possible while processing each text constituent.

Kintsch (1977) makes use of the story grammar approach to text analysis, which has developed out of ideas such as those of Rumelhart (1975). This approach uses sets of syntactic rules, constructed by analogy with rules of sentence syntax, which operate on symbols representing text-level categories. The structural descriptions which story grammars assign to stories are said to represent the abstract structural organisation of the text, and consist in essence of

hierarchical trees composed of different types of story constituent. For example, a crucial constituent in most story grammars is the Episode; the constituents of Episodes vary in different grammars, but generally consist of such categories as Goal, Attempt, and Outcome. An Episode is effectively a problem-solving structure in which the main protagonist of the story formulates some objective or desire and then goes through steps intended to achieve this aim. A story may consist of a series of Episodes, or may contain several Episodes embedded within each other. According to the Kintsch model, the types of information reduction which are permissible during macro-structure formation are constrained by the nature of the current story constituent which is being processed; macro-rules are seen as operating on each story category as a unit, to provide it with a summarising macro-proposition which captures the type of information which that category is intended to supply. For example, van Dijk (1977a) points out that readers who are processing the initial Setting of a story will know that state or process information will be more important than action or event information; the converse will apply while processing a Resolution category.

Kintsch (1977) presents a number of results demonstrating the role of such text analyses in macro-structure formation, but evidence for the psychological reality of story grammars is more plentiful outside the framework of the Kintsch model itself. One common finding (e.g. Mandler, 1978; Thorndyke, 1977) is that subjects' recall protocols of stories which are poorly structured in terms of the story grammar analysis tend to distort the text into a structure which is closer to the ideal story format; well-structured stories tend to be recalled more accurately than poorly structured stories. There is also evidence (e.g. Mandler and Johnson, 1977) that certain story categories are

recalled better than others; this clearly shows that readers differentiate between story constituents in memory. More recent research has shown that abstract story structure analyses also have observable effects during comprehension. Haberlandt (1980) found that subjects had longer reading times for constituents occurring at the beginning or end of an Episode than for constituents in the middle of an Episode; these large reading times were attributed to processes of macro-structure formation acting to summarise story categories. In a similar way, Mandler and Goodman (1982) found that the first sentences of constituents of Episodes had longer reading times than the second sentences; it was argued that the end of one constituent could only be detected at the beginning of the next, and that this was accompanied by summarising macro-processes.

There is thus considerable evidence for the psychological reality of story structures, and for their relevance to macro-structure formation. Although the story grammar approach has been criticised on various grounds, such as the formal inadequacy of the grammars used (Black and Wilensky, 1979), as Mandler and Goodman (1982) point out the validity of the structural analysis is not dependent on the validity of the formalism used to represent it. Texts undoubtedly have higher-level structures which play a role in comprehension and memory; the correct analysis of these structures is an empirical issue, and the value of the approach is not reduced by disagreements in the particular analyses, such as of the constituents of episodes, that are used by different researchers. The relevance of text structure to pronoun interpretation is brought out in the following section.

The Nature of Pronominal Reference

Pronouns are a closed class of short and relatively contentless lexical items which convey little inherent information. They are distinctive in that they cannot be fully understood independently, but must be interpreted relative to other components of the communication itself or of the situational context. The primary function of pronouns is therefore not to add new facts about an entity, although in certain cases they are able to do so, but to stand in place of or to represent concepts which are already known. The study of pronominal reference is largely a matter of determining what types of concept can be so represented, and of discovering the restrictions on which communicative components a particular pronoun can be related to.

One important distinction is between the deictic use of pronouns and the anaphoric use. Deictic pronouns must be related to some feature of the communicative context which is being in some way indicated or pointed to; anaphoric pronouns, however, are interpreted relative to some other component of the communication itself. In the case of personal pronouns, the first person forms such as 'I' and 'you' have only the deictic use (Lyons, 1968), and must be seen as representing the speaker and hearer respectively; third person forms such as 'he' or 'she' are however not marked as either deictic or anaphoric, and can be used in either way. Hence, pronoun interpretation is complicated by a logically prior difficulty of determining which form of relative interpretation is appropriate in any particular case. In reading, however, the deictic use is comparatively rare, and for present purposes

this difficulty will be ignored.

Anaphoric personal pronouns such as 'he' or 'she' typically stand for unique concepts mentioned earlier in the text; this can be distinguished from the cataphoric use of pronouns, in which the represented concept occurs later in the text. As the term implies, the antecedent element represented by a pronoun must generally be a noun phrase. Thus, in sentence (1), the pronoun 'he' clearly stands in place of 'the policeman'.

(1) The policeman liked whisky but he could only afford beer

It is worth noting that there are many other types of pro-form in English, which may stand in place of elements other than noun phrases. The pronoun 'it', for example, must on some occasions be related to entire propositions rather than to single noun phrases; this can be seen in sentence (2).

(2) The car has broken down but it doesn't bother me

Here, the pronoun cannot be interpreted relative solely to the noun phrase 'the car', but clearly stands for the entire stated fact of the first clause. Other types of pro-form may represent different types of information. The expressions 'do' or 'do so' may function as pro-verbs; while the expression 'such' may function as a pro-modifier, and stand in place of earlier qualifying adjectives. There may also be variations in the amount of information which a pro-form represents; a single anaphoric expression such as 'this' or 'so' may stand for a long and complex section of the preceding text. Examples of these different types of pro-form are given in de Beaugrande (1980).

In addition to different varieties of pro-form, English also possesses other ways of anaphorically representing noun phrases. The two most important of these are elliptical or deleted anaphors and the use of definite noun phrases. Ellipsis occurs when a syntactic element is omitted from the surface sentence but is inferred during comprehension; thus, in sentence (3) it is evident that the missing subject of the second clause is 'Harry'.

(3) Harry felt tired and went to bed

Chomsky (1981) points out that the reality of deleted anaphors is also attested by consequences for the phonetic representation of a sentence. The anaphoric use of definite noun phrases is only one of the functions of definiteness, but is illustrated by (4); here, the expression 'the vessel' must clearly be related to the earlier mention of 'a ship'.

(4) A ship sailed into harbour

The vessel docked smoothly

Given the existence of these different types of anaphora, it is perhaps curious that psychological research has concentrated almost entirely on the personal pronouns and on definite noun phrases. Although these expressions are undoubtedly important and interesting in their own right, the overall aim of studies of text processing is not to understand pronoun interpretation in isolation but to relate it to text integration and comprehension as a whole. As will be shown below, pronoun interpretation has on occasion been used as a tool to investigate other, more general features of text processing. It is possible that the neglect of other forms of anaphoric connection may lead to mistaken conceptions of reading comprehension, if in fact the interpretation of nominal anaphors is not fully representative.

It is clear that there are differences even among the several types of nominal anaphor. For example, de Beaugrande and Dressler (1981) point out that anaphoric subject deletion is not possible in a subordinate clause, while anaphora by means of a pronoun encounters no such restriction. Thus, the sentence (5) is unacceptable, while the same sentence with 'he' as the stated subject of the second clause would be well formed.

(5) Harry went to bed because felt tired

An important difference between anaphoric pronouns and anaphoric definite noun phrases will be frequently mentioned in what follows. While pronouns contain very little intrinsic semantic content, definite noun phrases can be constructed in order to be highly informative. The use of qualifying adjectives or relative clauses can render a noun phrase unambiguously specific, while even a single noun can provide important clues regarding the nature of its antecedent. These differences may have important psychological consequences.

One important issue has so far been glossed over here. It has been stated that pronouns stand in place of or represent concepts introduced earlier in the text; however, these expressions clearly say little concerning the precise nature of the relationship between pronoun and antecedent. The most common approach to this relationship is to say that pronouns and their antecedents are coreferential. Thus, a pronoun is not interpreted solely in relation to the text element which stands as its antecedent, but in relation to the referent of that element. However, pronouns and their antecedents need not in fact be coreferential. For example, Webber (1980) discusses examples of anaphora in which the connection between anaphor and antecedent is one

of shared sense, rather than of shared reference. Another form of anaphoric relationship which poses difficulties for the coreferential view is illustrated in (6). Here, the expressions 'a horse' and 'it' do not identify an actual animal, nor even an animal existing in the imaginary world of the sentence; they rather describe a hypothetical state of affairs in relation to the sentence itself.

(6) If John had a horse he would ride it every weekend

It therefore appears that anaphoric expressions may be interpreted relative to the sense of an earlier element rather than its reference, or relative to a state of affairs which is hypothetical even from the point of view of the text itself, as in (6). These difficulties are not avoided by the approach suggested by Webber (1980), who mentions the above problems and argues that pronoun referents should be seen as entities in the speaker's model of the discourse; in cases of identity of sense anaphora the pronoun would still require that a new model entity be constructed, while in (6) the sentence would need to be seen as leading to the construction of a model within a model.

For present purposes, these complexities in the nature of the anaphoric relationship will be ignored. Like the bulk of the psychological studies on pronominal reference, this thesis will be concerned solely with cases where pronoun and antecedent can satisfactorily be considered as coreferential within the world described by the text. However, just as the neglect of other types of pro-form leaves the generality of conclusions based on pronominal coreference open to question, so the neglect of other types of pronominal relations is also ultimately unsatisfactory. Since pronouns can be used in so many different ways, and since other types of pro-form must be related

to different kinds of antecedent, readers must have ways of identifying the appropriate comprehension strategies in each particular case. To treat one form of pronoun interpretation in abstraction from other types may therefore overlook a prior feature of comprehension with possibly significant repercussions for the aspects of anaphora resolution which a specific theory seeks to explain. This danger can only be noted here; yet it is evident that the study of coreferential anaphoric pronouns must eventually be united with other aspects of text integration.

Influences on Pronoun Assignment

The different types of pronominal anaphor do share at least one feature in common. However the relationship between pronoun and antecedent is conceived, it is evident that the correct antecedent must be identified in each case; whether the antecedent is used to construct a representation of a new entity based on its sense, or is related to some hypothetical case, or is interpreted as straightforwardly coreferential with the pronoun, successful interpretation requires that the antecedent used for these purposes be the one intended by the speaker or writer. The importance of this problem is reflected in the emphasis placed upon it by both linguistic and psychological research. Such research has attempted principally to identify the constraints which determine potential connection between particular pronouns and antecedents, or in the case of psychology to determine the processing heuristics by which readers actually choose antecedents. Other aspects of pronoun interpretation, such as the integration of information which must follow successful assignment, have received relatively little attention.

The most obvious constraints on pronoun assignment are derivable from the content of the pronoun itself. Although they convey very little inherent information, pronouns do at least place restrictions on certain general properties of the required antecedent. One of these is gender; in English, pronouns can specify whether the intended antecedent is male or female, as illustrated by the difference between 'he' and 'she'. A second restriction is that a pronoun may specify whether the antecedent is a single entity or a group of entities, as seen in the difference between 'he' and 'they'. Such restrictions can readily exclude many of the potential antecedents in a text, and in cases where only one element matches the gender and number specification of the pronoun they can fully determine the correct assignment. This is illustrated in (7), where gender cues allow the reader to determine that the antecedent of 'she' is 'Mary'.

(7) John played tennis with Mary

She won by two sets

It is not necessary that the antecedent to a pronoun fully matches its gender specification. In cases where it does not, the assignment may nonetheless proceed, but the pronoun specification is taken as adding more information about the entity involved. This is seen in (8), where it is only in the second sentence that the sex of the cyclist becomes apparent. Lyons (1968) points out that this is possible in English, since gender is largely 'natural' gender; in other languages, gender serves grammatical functions which do not require that the referent be actually male or female.

(8) A cyclist had an accident near the traffic lights

She was knocked down by a lorry

The use of gender cues in pronoun assignment has not received much psychological study in its own right, presumably because of its self-evidence; however, later it will be shown that gender cues can be usefully manipulated when investigating other issues. It should also be noted that the linguistic fact that gender cues may dictate an assignment does not establish the role of such cues in the actual processes of interpretation used by readers.

The fact that gender cues alone are insufficient to explain pronoun assignment is illustrated by (9) and (10). In both these sentences, two acceptable male entities are available as potential antecedents for 'him'; however, in (9) it is clear that the antecedent must be 'Fred', while in (10) neither of the two entities can be correctly taken as antecedent.

(9) Fred wants John to like him

(10) Fred wants him to be liked by John

The factors which determine that the pronoun can be related to 'Fred' in (9) but must refer to some third party in (10) are evidently syntactic in origin. In general, these examples show that potential coreference between a pronoun and an antecedent cannot be assessed by considering the two elements in isolation, but only by taking into account the structural context within which they occur. One linguistic explanation of the particular restrictions illustrated in (9) and (10) is offered by Chomsky (1981); however, the details of the syntactic analysis are not necessary for an appreciation of the fact that within-sentence anaphora is constrained by structural variables. Many such constraints have been noted in linguistics.

Another influence on assignment based in the structural relationship between pronoun and antecedent has been noted by Maratsos (1973) and Sheldon (1974). This is illustrated by sentences (11) and (12).

(11) Susan telephoned Mary and then she wrote to Jane

(12) Peter kicked Dave and then George punched him

The striking feature of these sentences is the tendency to assign the pronoun in the second clause to the individual who plays the same role in the first clause. Thus, in (11) 'she' is normally assigned to 'Susan', since both are in subject position; while in (12) 'him' is assigned to 'Dave', since both are in object position. This influence on assignment was termed parallel function by Sheldon (1974). Maratsos (1973) suggests that this is a simple and developmentally prior cognitive strategy which determines the natural pronoun assignment; variations in the strategy, to be described below, only appear in children's comprehension at a later age. Chafe (1976) points out that the preservation of roles can also operate across separate sentences; this is illustrated by (13), in which 'he' is normally assigned to 'Ted' and 'him' to 'Harry'.

(13) Ted saw Harry yesterday

He told him about the meeting

Grober, Beardsley and Caramazza (1978) made an explicit test of the parallel function strategy. They presented subjects with sentence fragments such as 'John may scold Bill because he...', and asked them to write a completion for each sentence which expressed a reason or motive

for the action described. These sentence completions were rated according to whether they indicated that the pronoun had been assigned to the first noun or to the second. It was found that there was a strong tendency for the pronoun to be assigned to the first noun in the sentence, which clearly played the same subject role as the pronoun itself.

The parallel function strategy therefore appears to be genuine. There are, however, two points to be made. Firstly, as Maratsos (1973) points out, it is not clear whether the strategy seeks to preserve the syntactic or the semantic roles of the elements in a sentence; in the examples given above, grammatical subjects also play the role of actor or agent, while objects play the role of recipient or experiencer. Secondly, unlike the structural influences exemplified by (9) and (10), role preservation can at best be described as a tendency; parallel function does not rigidly dictate the assignment of the pronoun. This is illustrated by other results of Grober et al. (1978), who found that parallel function assignments to the sentence subject could be overridden if semantic features of the verb indicated that it was the second person mentioned who would be most likely to have caused the stated action. This influence on assignment has been termed implicit causality, and is discussed in more detail below. In general, the majority of assignment influences which have been studied in psychology must be considered in a similar fashion; rather than determining an assignment, the influences simply indicate one which is most probable.

The parallel function strategy can also be overridden, in spoken sentences at least, with appropriate use of stress. As pointed out by Akmajian and Jackendoff (1970), if for example sentence (11) were spoken

with heavy stress on 'she', the pronoun would tend to be assigned to, 'Mary' rather than 'Susan'. Maratsos (1973) studied the interpretation of the stressed and unstressed forms in children aged from 3 to 5. When the children were asked to use small dolls to act out events described in such sentences as 'Susie jumped over the old woman and then Harry jumped over her', children of all ages showed that they had correctly assigned the pronoun to the entity fulfilling the same role in the first clause if the pronoun was unstressed; when the pronoun was stressed, however, the younger children were found to be less able to make the reversal in assignment which such pronunciation indicates. Maratsos hypothesised that stress acted as a signal that the natural interpretation of the sentence was being breached.

A similar effect of stress on pronoun interpretation can be seen in the phenomenon of exhaustive listing. Kuno (cited by Webber, 1980) points out that while the entire group of individuals given in an exhaustive listing can be referred to pronominally, no one member of that group can be referred to by a pronoun. Thus, sentence (14) can be followed by (15), but not by (16).

(14) It was Mary, John and Susan who failed French 'O' level

(15) They also failed German

(16) He also failed German

The exhaustive listing constraint on pronoun assignment is interesting in its own right, since it demonstrates again that the context of occurrence of a prior element must be taken into account when determining whether it is a possible antecedent; although 'John' is explicitly available in the prior sentence, it cannot be anaphorically referred to by 'he' in (16). This constraint, however, can apparently

be overridden by pronouncing (16) with stress on the initial pronoun, in which case the coreference with 'John' becomes acceptable. Like the effect on parallel function, this outcome can be explained by regarding stress as an explicit signal that a normal influence on assignment is being suppressed. From a psychological point of view, an important question is whether the normal assignment in such cases is easier to process than the unusual stressed assignment. As will be discussed below, one theory of pronoun assignment argues that certain entities are inherently more available as antecedents than others; the fact that patterns of availability can be altered by stress may be some evidence against this view, but only if the different patterns of assignment are processed equally easily.

So far, this review has been concerned principally with syntactic or structural relationships between pronoun and antecedent, and the restrictions these place on possible or probable assignments. However, various semantic and pragmatic factors also influence assignment. One of these, implicit causality in verbs, has already been briefly mentioned. This notion was originally put forward by Garvey and Caramazza (1974), who argued that when certain verbs were used with animate nouns they carried an implicit attribution of the cause of the described action. Thus, the sentence 'the mother punished her daughter' may be taken to imply that the child had done something wrong which led to her being punished, rather than, for example, that the mother was simply a cruel parent.

It has already been mentioned that implicit causality can counteract the influence of parallel function. Thus, in (17) the pronoun is in subject position in the second clause, and so on the basis of parallel function would be taken to relate to 'John'; the evident

preferred assignment, however, is to 'Bill'. According to Garvey and his co-workers, this assignment derives from the intrinsic bias of the verb 'blame', which makes the object of the action the more likely source of the reason for the action.

(17) John blamed Bill because he spilled the coffee

The existence of such biases was demonstrated by Garvey, Caramazza and Yates (1976). Like Grober et al. (1978), Garvey et al. presented subjects with sentence fragments of the form 'NP1 V NP2 because PRO...' and rated the subjects' completions according to which NP was taken as the pronoun's antecedent. It was found not only that no subjects reported noticing the potential ambiguity of the pronoun, but also that most of the test items elicited clear assignment biases; verbs such as 'scold' or 'kill' led to sentence completions indicating assignment to the second NP of the first clause, while verbs such as 'sell' or 'confess to' led to completions indicating assignment to the first NP of the sentence. These biases were corroborated by subjects' responses to questions of the form 'Why did NP1 V NP2?'.

Although implicit causality therefore appears to bias assignment when subjects are presented with sentence fragments, verb features are not a rigid constraint on antecedent choice in full sentences. In particular, such biases can be overcome if the content of the second clause indicates that the biased assignment would be incorrect. Thus, while in (17) the content of the second clause is consistent with the biased assignment to 'Bill', in (18) the second clause contradicts the verbal bias in favour of 'Mary' and in fact makes 'Jane' the most plausible antecedent to 'she'. Gender cues may also override a verb bias, as in (19).

(18) Mary confessed to Jane because she was sympathetic

(19) The father punished the daughter because he was in a bad temper

Caramazza, Grober, Garvey and Yates (1977) investigated the effects of pairing a biased verb with an inconsistent second clause by presenting subjects with full sentences and asking them to say the name of the pronoun's referent out loud. Two important results were found. Firstly, referent choice latencies were faster when the sentences included an unequivocal gender cue to the correct antecedent than when they did not; extensive discussion will be given below to the view that this indicates that pronouns are initially assigned on the basis solely of their inherent content. Secondly, latencies were faster when the semantic bias from the content of the second clause, or the assignment constraint derived from gender cues, was consistent with the implicit bias of the main verb. It was concluded that subjects took account of verbal biases when choosing a referent for the pronoun, even when this was strictly unnecessary as a result of gender cues; and that overriding the verb bias by other factors caused comprehension difficulties.

The interesting feature of the above experiment is that it made an explicit test of interactions between different assignment influences, and apparently showed that verbal biases were pre-eminent factors which could only be overridden with difficulty. Grober et al. (1978) found that the verb bias could also take precedence over parallel function, but presented no latency measure of the consequences of such inconsistency for ease of processing.

Ehrlich (1980) challenged the view that referent choices are determined by verb biases in preference to pragmatic inferences based on sentence content. She argued that the important aspect of the results of Caramazza et al. (1977) was that subjects did in fact choose the pragmatically indicated referent in cases where the verb bias was inconsistent with this assignment, and claimed that readers make use of general knowledge of the described events when determining the pronoun's antecedent. She tested this, not by varying the content of the second clause, but by varying the linking conjunction. Thus, sentence (20) is a repetition of (17), with the conjunction 'but' in place of 'because'; it is clear that this indicates the correct antecedent of the pronoun to be 'John', not 'Bill'.

(20) John blamed Bill but he spilled the coffee

Ehrlich presented subjects with sentences of this type, containing either the conjunction 'because', 'but', or 'and', and required them to make referent choices by releasing a button and then saying the selected name out loud; half the sentences contained a gender cue dictating the assignment, while half did not. Several interesting results were obtained. Firstly, use of the conjunction 'but' was found to reverse referent choices compared to sentences containing 'because'; Ehrlich argued that this showed subjects were using knowledge of event relations to select a referent. Secondly, referent choice button release latencies were faster when a gender cue was present than when there was no gender cue; this was similar to the result of Caramazza et al. (1977). Thirdly, referent choices appeared to be slower when the conjunction 'but' was used than when 'because' was used; this outcome is similar to the finding of Caramazza et al., that inconsistent sentences led to greater processing difficulty, but on Ehrlich's method

of data analysis this did not prove fully significant.

Ehrlich's conclusion, that general knowledge constraints affect pronoun assignment, cannot seriously be controverted; however, this fact does not address the crucial psychological question. The main aim of the psychological study of pronoun assignment is not simply to identify the ultimate influences on referent selection, but to specify the processes which readers carry out when making such choices. To this extent, Ehrlich is wrong in saying that the important fact is that a certain antecedent is selected; the more important matter is how it was selected. Latency data are more informative on this issue than assignment probabilities. Thus, the fact that latencies are slower when pragmatic or general knowledge factors are in conflict with verbal biases, as was found both by Caramazza et al. (1977) and to some extent by Ehrlich herself, suggests that verb features are considered when selecting a referent even though they may not determine the final assignment. Ehrlich's results concerning preferred assignments with the conjunction 'but' had in fact already been anticipated by Grober et al. (1978). Some of the sentence fragments presented to subjects by these authors did contain 'but' in place of 'because', and it was found that completions clearly indicated choice of the first sentence NP as antecedent to the pronoun. As noted above, this was explained in terms of the parallel function strategy.

One of the latency results found by Ehrlich (1980) was that referent choices were faster in the presence of a gender cue; she argued that assignments were initially made on this basis if possible, with general knowledge being used only if there was no unequivocal gender constraint. However, in all Ehrlich's test sentences the gender cue was consistent with the pragmatic assignment bias based on event

relations; thus, faster assignments in the presence of a gender cue may have been due to agreement between two combined influences, rather than to initial preference for one over the other. The result of Caramazza et al. (1977), that latencies were slower when the gender cue conflicted with the assignment bias of the verb, supports the view that both sources of information were called on in making the decision. Hirst and Brill (1980) also showed that pragmatic influences on assignment were used even when a pronoun was clearly unambiguous. In one of their experiments, subjects were presented with sentence pairs such as (21) or (22). In both cases, the pronoun 'he' is syntactically constrained such that it must refer to 'John'; however, in (21) the content of the first clause of the second sentence is clearly more likely to relate to 'John' than to 'Henry', while in (22) this difference is less acute. These differences in semantic plausibility had been determined in a preliminary ratings study.

(21) John stood watching

He ran for a doctor after Henry fell down some stairs

(22) John stood watching

He thought of the future after Henry fell down some stairs

The subjects' task was again to make an explicit referent choice, and to press a button to indicate which individual had been selected. Despite the presence of a syntactic constraint, assignment latencies were found to be faster in cases with a clear plausibility cue. Hirst and Brill argued that this result indicates that information in the two sentences is integrated even during processes of pronoun assignment, rather than after assignment as is claimed by the given-new strategy; however, a more likely explanation is that readers made combined use of

syntactic and semantic cues when deciding on a referent for the pronoun.

In summary so far, it is evident that there are many different influences on pronoun assignment. Some of these, such as gender cues or syntactic constraints, can rigidly determine the antecedent for a particular pronoun; others, such as parallel function, implicit causality or pragmatic inferences, can only specify that a certain assignment is highly probable. Not only do factors of this second type interact, such that for example conflicts between implicit causality and pragmatic cues lead to difficulty in assignment, but also even influences of the first type can be hindered by other, contradictory assignment signals. Assignment choices are dependent on structural and semantic relationships between pronoun and antecedent, and are made on the basis of combined information of different types.

This view of pronoun assignment is not, however, universally accepted. One widely held theory argues that initial attempts at pronoun assignment are made on the basis of very limited information, and in particular using only the specification in terms of gender and number which is derivable from the pronoun itself. In addition, this theory argues that certain entities in a text are intrinsically more likely to be taken as the antecedent of a pronoun than others; assignment to these entities is seen as occurring without the use of complex relational or pragmatic cues. Ehrlich (1980) is one exponent of this view. It has already been noted that Ehrlich took her results to show that assignments were attempted on the basis of gender cues before pragmatic constraints were taken into account; she also argues, however, that her results indicate that sentence subjects are generally taken as antecedents in preference to objects. She found that her

experimental subjects tended to choose the verbally or pragmatically biased NP more frequently when it was the first noun in the sentence than when it was the second; that is, biases to the object of the first clause were relatively ineffective. This apparent indication of a preference to select the subject NP in spite of other biases is however cast in doubt by the fact that the crucial pronouns were themselves always the subject of the second clause. The obtained assignment probabilities may therefore be due to the parallel function strategy, rather than to the intrinsic status of sentence subjects. It may incidentally be noted here that no on-line processing studies of parallel function appear to have been reported in the literature; it is at least possible that parallel function assignments would prove to be easier when it was the subject role which was preserved across clauses rather than the object role.

The alleged preference for sentence subjects as antecedents rather than sentence objects has been asserted by other authors in addition to Ehrlich (1980). Levin, as described in a highly abbreviated listing of his suggested assignment heuristics given by Carpenter and Just (1977a), argues not only that sentence subjects are preferred to objects, but also that objects are preferred to antecedents in prepositional phrases. Sanford and Garrod (1981) also assert the preference for subject assignments, and describe an experiment in support of it. In this study, subjects read short passages presented sentence by sentence on a computer screen; each passage contained a target sentence which anaphorically referred to an earlier entity by means of either a pronoun or definite noun phrase. The crucial antecedent sentence was separated from the target by either one or three intervening sentences; and the antecedent itself occurred in either subject or object position. Several results of interest were found. Firstly, reading times for the

target sentences were faster with one intervening sentence than with three; this distance effect is returned to in the next section. Secondly, anaphoric references to antecedents in object position produced slower reading times than references to antecedents in subject position; this can be explained on the view that subjects are preferred antecedents, and that assignments to objects were hindered by the initial attempt to assign the pronouns to the subject. Thirdly, for subject antecedents target sentences containing pronouns were always read faster than those containing definite noun phrases, while for object antecedents this pattern was reversed when there were three intervening sentences. This difference can be explained by arguing that subject assignments are easier due to their intrinsically greater availability, while object assignments must be signalled with the explicit semantic cues of the definite noun phrase.

The interpretation of this experiment will shortly be returned to in a wider context, but for the present its relevance to the alleged preference for subject assignments is most important. Unfortunately, however, it appears that this study, like that of Ehrlich (1980), may have been contaminated by effects of parallel function. Sanford and Garrod (1981) give only two examples of the passages used, but in both it is striking that the crucial anaphor occurred in subject position in the target sentences. Clearly, this may have been one source of the greater ease of subject assignments, if the pattern was common in the other experimental materials.

The alleged preference for sentence subjects as pronominal antecedents is a special case of the view that aspects of linguistic structure may indicate that one particular entity is most likely to be referred to again. Another instance of such effects on antecedent

preferences was studied by Carpenter and Just (1977a). They argued that certain syntactic structures, illustrated in (23) and (24), single out one of their components as the preferred antecedent for subsequent pronouns. Sentence (23) is a cleft sentence; here, Carpenter and Just argued that 'the arsonist' is highlighted. In the pseudo-cleft sentence (24), 'the arsonist' is again highlighted, though in a somewhat different manner. According to Carpenter and Just, sentence (25) provides no such strong linguistic cue which could determine a preference between the potential antecedents.

(23) It was the arsonist who the guard mocked

(24) The one who the guard mocked was the arsonist

(25) The guard mocked one of the prisoners in the machine shop

Carpenter and Just presented subjects with short passages in which a pronoun sentence followed either a highlighting sentence such as (23) or (24), or a more neutral sentence such as (25). The pronouns were always ambiguous on lexical, syntactic, and pragmatic grounds, such that any of the preceding entities was a possible antecedent. Subjects read the passages while their eye fixations were recorded; their task was to decide whether each sentence was consistent with or contradicted what had gone before, and at the end of each passage to give prompted recall of as much as they could remember. It was found that when subjects encountered an ambiguous pronoun, they tended to make regressive eye movements to the preceding sentence; when this sentence was a cleft or pseudo-cleft, regressions were made most frequently to the highlighted entity, while if the sentence was more neutral the different entities were fixated approximately equally often. In the recall protocols, it

was found that subjects frequently replaced the critical pronoun with one of the earlier noun phrases; this was presumably the one which had been taken as the pronoun's antecedent. When the preceding sentence had highlighted a particular entity, it was that noun phrase which was most frequently substituted for the pronoun.

Carpenter and Just (1977a) have thus demonstrated one important influence on assignment. In the absence of determining cues such as gender or syntactic constraints, and without assistance from pragmatic disambiguation, subjects nonetheless showed an apparent preference for one potential antecedent over the other; this preference was based on sentence structure cues. These preferences were not however overwhelming, since both the regression and the recall data indicated that on roughly a third of occasions subjects selected the non-highlighted entity as the pronominal antecedent. Furthermore, it must be stated again that the demonstration of an influence on assignment does not reveal the way in which that influence produces its effects. It is possible, for example, that readers fall back on predetermined antecedent preferences only when other sources of clear disambiguation are not available; alternatively, they may attempt to combine all sources of assignment influence simultaneously, with the eventual referent choice depending on the outcome of this evaluation. Finally, the Carpenter and Just results only demonstrate an assignment preference; they do not indicate whether assignments to highlighted entities are achieved more rapidly or easily than those to less clearly marked entities.

An experiment very similar in intention to that of Carpenter and Just (1977a) was carried out by Yekovich, Walker and Blackman (1979); these authors were however explicitly concerned with the effects of

antecedent marking on ease of anaphor assignment, rather than on preferred assignments. They hypothesised that antecedent nouns should ideally be marked as new information in the sentence in which they occur, while anaphors should be marked as old information; their study was in fact concerned with definite noun phrase anaphora rather than with pronouns, and it is evident that definite noun phrases need not always represent anaphoric, given information. Subjects were presented with sentence pairs, in which a noun phrase from the first sentence was anaphorically repeated in the second sentence; both the antecedent and anaphor could appear in sentence positions where they were marked as either given or new. The sentence pair in (26) illustrates what Yekovich et al. considered to be the ideal marking of antecedent and anaphor. Here, 'the shark' occurs as object of the main clause in the first sentence, and hence is marked as new, but occurs as subject of the main clause in the second sentence, and hence is marked as presupposed or given. Evidently, the suggestion that sentence object position is the ideal location for an antecedent contradicts the view that sentence subjects are taken as preferred antecedents, although this contradiction may possibly be attributable to differences between pronominal and definite noun phrase anaphora.

- (26) The lifeguard spotted the shark from a tower on the shore
The shark attacked the diver near the reef

Reading times for second sentences presented with the ideal marking illustrated in (26) were compared with conditions where either the antecedent was marked as presupposed, the anaphor marked as new, or both. Reading times were found to be fastest in the ideal case, slowest when both antecedent and anaphor were incorrectly marked, and intermediate in the other two conditions. Yekovich et al. concluded

that the various entities mentioned in a sentence are not all equally easily referred to again, but that certain linguistically marked elements are more accessible to anaphora. In processing terms, they suggested that information which is linguistically marked as new would receive greater strength of coding in memory, and hence a search of memory cued by the anaphoric term could locate its antecedent more readily.

The demonstration by Yekovich, Walker and Blackman (1979) of differential ease of anaphoric reference to antecedents in different syntactic positions is an example of a central finding in the psychological study of anaphora. Many studies have shown that some antecedents may be intrinsically easier to refer to than others, and have related this difference to the effects of signals provided by the text. This finding has been fundamental in the construction of processing models of pronoun assignment, and forms the main concern of the following section.

The preceding section gave a review of many influences and constraints on pronoun assignment. Some of these influences were based on lexical and syntactic factors which can fully determine particular assignments, while others consisted of more flexible tendencies or preferences to select certain antecedents in particular cases. However, a simple description of the various influences on pronoun assignment is insufficient for psychological purposes. The principal goal of psychological research on pronoun interpretation is not merely to identify the different influences which may be at work, but to devise a processing account which explains how these influences operate and interact, and which relates them to other aspects of comprehension. Although the influences described earlier took several different forms, a meaningful psychological theory must nonetheless be able to accommodate them all within a single framework. This framework must specify how and when readers draw on the different sources of information which can potentially influence assignment, and must show how these influences eventually lead to correct antecedent selection.

One way of reducing the large number of apparently distinct influences to more manageable proportions is to attempt to identify similarities between them, such that several diverse factors can be treated in a single explanation. It is evident that some of the last-mentioned influences in the preceding section can be combined in this way. It was noted there that items were allegedly more likely to be selected as pronoun antecedents, or were selected more rapidly, if they appeared in sentence subject position, were highlighted by a cleft construction, or were part of the linguistically marked new information;

this pattern of effects suggests that certain text entities receive preferential status with regard to pronoun interpretation. In the absence of semantic or pragmatic factors which select one referent among several potential referents, readers nonetheless show clear preferences for certain marked entities (Carpenter and Just, 1977a); and pragmatic biases themselves can apparently be weakened by conflicting assignment preferences based on antecedent status (Ehrlich, 1980). It therefore appears that some entities can be made intrinsically easier to refer to by variations in the way they are mentioned in the text. This possibility can be generalised still further, by saying that in appropriate circumstances some text entities may be inherently easier to refer to than others; this generalisation omits the condition, based on the results listed above, that such ease of reference must be due to the manner of the initial mention of an entity. The suggestion being made is simply that the independent status of an item, however this may be determined, can be a crucial factor governing the ease with which it can be referred to pronominally.

The fact that certain entities in a text may be intrinsically easier to refer to than others was noted by Chafe (1972). The feature of antecedent status which governs accessibility to pronominal reference was there termed foregrounding, and this has become the standard label for this phenomenon in the psychological literature. According to Chafe, antecedent foregrounding is a necessary but not sufficient condition for pronominal reference; foregrounded items may also be referred to with definite noun phrases, but the use of a definite noun phrase does not require that the antecedent be foregrounded. A reflection of foregrounded status is also found in speech, where references to foregrounded entities are pronounced with low pitch except in cases of contrastive emphasis. In Chafe's view, foregrounded

concepts can be weakly pronounced because "...they are already in the minds of the participants in the discourse, who are able to reserve louder and higher pitched pronunciation for the elements in the sentence which provide new information." (Chafe, 1972, p51). Pronunciation of foregrounded items with weak stress is thus similar to the observation by Halliday (1970) that given information is pronounced with weak stress; however, Chafe (1972) is in fact arguing that only certain readily available given concepts can be so pronounced. As will shortly be indicated, Chafe's conception of foregrounded items as being present in consciousness has been translated directly into psychological terminology by many researchers.

Chafe's paper did not view foregrounding as something specifically determined by the manner in which an entity was first mentioned, but concentrated instead on variations in foregrounded status which took place in the text following an item's first occurrence. Chafe argued that foregrounding would gradually evaporate if an entity ceased to be mentioned in succeeding sentences, and that foregrounded status would also be lost due to changes of scene in the text, or due to the introduction of new topics. Items could be considered foregrounded immediately after their first mention, but eventually would lose this status and hence would no longer be accessible to pronominal reference. The possible regaining of foregrounded status was discussed by Lewis (1979), who in fact referred to such status with the term salience. Lewis pointed out that in the course of a single monologue a simple noun phrase, such as 'the cat', can be used to refer to quite distinct entities; the choice of referent in each case would depend on the relative salience of the available alternatives. Reintroduction of an entity low in salience was seen as increasing that entity's salience level once more.

The linking factor between these observations concerning foregrounding in the text following an entity's first mention, and the assignment influences noted above that were based on the manner of that first mention, is the concept of topicality. Carpenter and Just (1977a) made this connection explicit by referring to the information highlighted in cleft constructions as the foregrounded information, and by claiming that this information was indicated by a topic marker which they termed the discourse pointer. In Carpenter and Just's view, the current topic of discourse is taken as the initial preferred antecedent for anaphoric pronouns, and readers maintain an internal discourse pointer indicating that topic; ease of reference to an entity is thus dependent on its topical status. If topical status is seen as being lost during succeeding sentences in which the entity is not mentioned, and as being regained if that entity reappears in the text, the observations of Chafe and Lewis can be unified with the experimental results already described.

The statement that foregrounded entities are the initial preferred antecedents to anaphoric pronouns is not in itself a complete psychological processing account. In order to provide a fuller specification of the ways in which foregrounding operates, most of the researchers whose work is referred to below have made use of the distinction between long-term memory and working memory. As was described earlier, working memory is the most common term used for a shared central processor which possesses limited, short-duration storage capacity; in reading, working memory is seen especially as playing a role in sentence comprehension, since it can maintain early parts of a sentence for connection with later input. The contents of working memory are distinct from the long-term memory representation of the text, which can be more complex and contain much more information. To

relate these distinctions to pronoun assignment, it has been widely argued that working memory is used to hold representations of the currently topical or foregrounded entity; information in working memory can by definition be directly accessed by comprehension processes, while information present only in the long-term text representation is relatively difficult to access. Thus, if when a pronoun is encountered the normal processing strategy is to attempt to assign it to some entity in working memory, anaphoric references to foregrounded entities will necessarily be easier than those to backgrounded entities. Backgrounded entities could only be located in the long-term representation, which would entail additional processing efforts after attempted matches to working memory had failed. These suggestions are very close in spirit to the differences between immediate matches and reinstatement matches proposed by Kintsch and van Dijk (1978). It is evident, and was illustrated by the importance to the Kintsch model of the nature of the buffer selection strategy, that this approach places the emphasis on the predetermined status of intended antecedents as the main factor governing the ease of pronoun interpretation. Pronoun interpretation will be easy if the intended antecedent is present in working memory, but will be difficult if the antecedent is present only in the long-term text representation. In addition, ambiguous pronouns will tend to be assigned to foregrounded entities rather than to backgrounded entities. Since the status of antecedents is seen as the crucial issue in pronoun interpretation, this approach has led to a research emphasis aimed at identifying the principal textual determinants of foregrounded status. The main strands of this research will now be reviewed.

One influence on foregrounded status suggested by Chafe (1972) was textual distance; the more successive sentences there were which did not mention a concept, the less likely it would be that that concept

would remain foregrounded. On this reasoning, pronouns would be expected to be more difficult to interpret the greater the textual separation between them and their antecedents. When pronoun-antecedent distance was small, the presence of the foregrounded antecedent in working memory would allow rapid assignment; when this distance was large, assignment to the now backgrounded antecedent in long-term memory would be relatively difficult. Chafe suggested that whether an entity was foregrounded or backgrounded might be uncertain at intermediate distances, but the distinction between near and far cases is clear.

The distance effect on pronoun interpretation has been frequently studied. Clark and Sengul (1979) contrasted two different models of the effect. According to the continuity model, the relationship between distance and ease of pronoun interpretation should be linear; the most recently mentioned entity should be most easily referred to, while entities mentioned progressively further back should be increasingly difficult to refer to. On the discontinuity model, however, a privileged place in working memory would be given only to the contents of the most recent sentence or clause; entities mentioned further back in the text than that privileged unit would be equally difficult to refer to irrespective of surface distance. Clearly, the former model would be compatible with the intuitive suggestions of Chafe (1972), while the latter model would not. To test these two models, Clark and Sengul presented subjects with four-sentence paragraphs in which the final sentence referred back to an earlier entity by means of either a pronoun or a definite noun phrase. The three context sentences were presented together, followed by separate display of the anaphoric sentence alone; this allowed recording of target sentence reading times. The crucial antecedent entity was mentioned in either the first, second or third context sentence, so varying distance from the anaphor.

It was found that target reading times were fastest when the antecedent was in the immediately preceding sentence, and equally slow if it was in either of the two earlier sentences; the results were identical for the two types of anaphor. In two further experiments using only definite noun phrase anaphors, Clark and Sengul (1979) showed that target reading times were facilitated only if the antecedent occurred in the last clause of the immediately preceding sentence; reading times were slowed if the antecedent occurred in the first clause of the preceding sentence. It was concluded that the evidence favoured the discontinuity model rather than the continuity model, and that only the most recent clause was privileged in working memory.

Clark and Sengul's results therefore contradict the intuitive proposal by Chafe (1972), who argued that foregrounding would evaporate gradually with distance. It appears instead that foregrounded status is lost immediately if an entity is not mentioned in the clause following its first mention. Unfortunately, the conclusion that additional pronoun-antecedent distance produces no further effects on ease of pronoun interpretation has not received published replication; the majority of other studies investigating the distance effect have used only two distance conditions, with the pronoun either adjacent to or separated from its antecedent. This procedure clearly does not reveal whether increases in interpretation difficulty are linear or discontinuous. One example of such a study is that of Cirilo (1981), who was concerned with definite noun phrase anaphora. Subjects in this study read short stories in which the anaphoric target sentence either immediately followed the sentence containing the antecedent or was separated from it by three intervening sentences. The passages were read under three different instructional conditions intended to make subjects emphasise either micro-processes of comprehension,

macro-processes of comprehension, or both equally. Only in the instructional condition intended to call up both types of processing, and which was therefore most equivalent to normal reading, was it found that target reading times were longer in the separated case than in the adjacent case. In the condition designed to emphasise macro-processing, which was expected to make readers less likely to build up specific referential connections between sentences, the distance effect did not appear. This confirms the view that the effect is caused by difficulty in locating antecedents to anaphoric expressions. The lack of a distance effect in the micro-processing condition was explained on the grounds that that condition involved subjects in giving immediate recall of small text fragments, and hence produced strategies different to normal reading.

There is thus evidence that self-paced reading times for anaphoric target sentences are affected by distance from the antecedent; this is true for definite noun phrases as well as for pronouns, and is explicable by arguing that antecedents are present in working memory only in the near conditions. However, this distance effect has also been demonstrated with a quite different technique by Daneman and Carpenter (1980). These authors tested ease of pronoun assignment by presenting subjects with comprehension questions after each experimental passage which could only be answered correctly if the pronoun had been assigned to the intended referent. This is illustrated in sequence (27), which consists of the final three sentences from one of the example passages which Daneman and Carpenter (1980) provide. The first sentence of the sequence contains the initial mention of a leopard; this crucial entity is referred to by a pronoun in the final sentence of the passage. After reading the passage, subjects received the question "Who finally arrived?". According to Daneman and Carpenter, such

questions as this would be more difficult to answer the greater the separation between pronoun and antecedent.

(27) The proceedings were delayed because the leopard had not shown up yet

There was much speculation as to the reasons for the midnight alarm

Finally he arrived and the meeting could commence

Daneman and Carpenter presented subjects with passages of this type in which pronoun and antecedent were separated by from one to six intervening sentences; the passages were presented in full on typed sheets, rather than on a computer screen as is usual in reading time studies. Unlike most studies, which ensure that all experimental passages can be varied to fit into all conditions in order to minimise the risk of specific materials effects, Daneman and Carpenter presented the same passages in each distance condition to all subjects. This was done in order to allow subjects to be divided into comparable groups on the basis of their reading ability, after the experiment was completed; this aspect of the experiment will be discussed in full in the next section. The overall results concerning the distance effect itself showed pronoun question accuracy to be significantly greater at shorter distances. This result could be readily explained by saying that referential connections between pronoun and antecedent could be more readily computed when the antecedent was in working memory at the shorter distances. However, the distance effect was not discontinuous; pronoun question accuracy continued to decline as distance increased, rather than showing a single fall.

There is thus considerable evidence indicating that increased distance between pronoun and antecedent leads to added processing

difficulty; other studies reporting such an effect, though in too little detail to be discussed here, are Dutka (1980) and Frederiksen (1981). However, there is some disagreement concerning the nature of the distance effect, and the nature of the working memory processes which may be responsible for it. Clark and Sengul (1979) argued that only the most recent clause was privileged in working memory, and their reading time data indicated that additional distance did not lead to further increases in processing difficulty; this argument is consistent with the view that working memory contents are cleared and replaced under the control of selection strategies which are sensitive to syntactic structure. However, Daneman and Carpenter (1980) found that pronoun question accuracy continued to decline as distance was increased, and evidently regarded antecedent presence in working memory as dependent on probabilities derivable from working memory capacity. Although he produced no data in support of this argument, Cirilo (1981) also stated that the probability that an antecedent would no longer be in working memory would be an increasing function of distance. This view therefore regards working memory contents in reading as dependent on general capacity and duration limitations, with items being gradually lost by processes of decay and displacement.

One way of reconciling the Daneman and Carpenter results with the discontinuity model of Clark and Sengul (1979) is through criticism of the technique which Daneman and Carpenter used. While target sentence reading time is a direct, on-line reflection of processing difficulty, responses to questions presented after reading a text may clearly be subject to other influences. In particular, question accuracy may be partly a function of inference processes occurring at the time of answering, rather than being a direct reflection of what has already been stored in memory. This can be seen in sequence (27). While the

question "Who finally arrived?" could in one way be answered on the basis of the final sentence of the sequence, assuming the pronoun had been correctly assigned, it could also be answered simply by recalling that the leopard was the one who had not yet shown up. Such inferences as these would not be dependent on correct assignment of the pronoun, yet might be sensitive to the distance back in the passage of the antecedent information on which the inference was based. It could therefore be argued that absence of the antecedent from working memory hampered pronoun assignment to a fixed extent irrespective of distance, but that pronoun question accuracy continued to decline owing to the increasing difficulty of making the necessary inferences.

The pronoun question technique is therefore ideal for testing neither the difficulty of pronoun assignment nor the issue of whether a pronoun has in fact been successfully assigned at all. However, it must be pointed out that the experiment reported by Sanford and Garrod (1981), and described in the previous section, did find some evidence of an increasing distance effect on reading times. That experiment found target reading times to be longer with three intervening sentences than with only one intervening sentence; this effect should not have occurred if non-adjacent antecedents were uniformly difficult to locate. One aim of the first set of experiments in the present thesis was to clarify the nature of the distance effect, and discover whether it was best regarded as discontinuous or linear. The decision between models suggesting active strategic control of working memory contents, and models arguing for gradual and probabilistic loss of working memory contents, is however also affected by consideration of the other known influences on foregrounding.

In addition to distance, the other influences which Chafe (1972) suggested might affect foregrounding were changes of scene and the introduction of new events. If a text moves on to describe actions or events not involving a certain entity, then it is plausible to assume that that entity's lack of relevance to the current text will cause it to lose foregrounded status. As with distance, Chafe suggested that the effects of such topic boundaries may be gradual and increasing; the greater the number of intervening events or changes of scene, the more likely it would be that a given entity would cease to be foregrounded. Evidently, distance and topic changes are inherently confounded. The greater the distance since an entity was last mentioned, the greater the likelihood that the current text will be concerned with topics unconnected to that entity; and conversely, if the current text is not related to a certain entity than that entity will necessarily go unmentioned. This raises the question of whether obtained distance effects reflect an independent influence on pronoun interpretation, or whether distance is simply an indirect measure of topic changes.

Sanford and Garrod (1981) describe an interesting experiment, conducted by Anderson, which shows how changes of scene may reduce foregrounded status. Sanford and Garrod regard comprehension as involving the use of scenario representations of stereotypical situations or events; the concept of a scenario is very similar to that of a script (Schank and Abelson, 1977). If a text describes events or actions taking place in a standard situation, scenario information will be activated specifying the most likely individuals that will be encountered in that situation, as well as details on the time and space constraints which that situation is normally subject to. Thus, a scenario for having a haircut would specify the existence of an assistant to cut the hair, and would stipulate a normal time range

within which the visit to the barber's shop would probably be completed. It was hypothesised that, if descriptions of scenario-based events were followed by a sentence stating an explicit time change, such as 'Five hours later...', then continued availability of scenario characters would depend on whether or not that time change exceeded the scenario's normal limits. When the time change exceeded those limits, entities taking part in the scenario would cease to be foregrounded, and hence would be difficult to refer to pronominally. This is clearly a refinement of Chafe's notion of changes of scene, with readers considered to be making use of stereotyped knowledge to infer a scene change even where it is not made explicit in the text.

An experiment was conducted to test the effects of time changes within and beyond normal scenario limits. Passages were read in which target sentences containing pronominal references to scenario-dependent entities were separated from descriptions of the scenario events by a sentence stating a large or small time change; target sentence reading times were recorded, as were response latencies to subsequent questions concerning the antecedent entity. It was found that both target reading times and question response latencies were slower after a time change exceeding the scenario limits than after a time change remaining within those limits. Interestingly, reading times for the sentences stating the time change were longer for large changes than for small changes; since the same sentences were used in each case, with only one word being altered, this is a striking effect. Sanford and Garrod interpret the results as demonstrating that scenario-based entities are cleared from working memory if textual cues indicate that they are no longer relevant, and attribute the effect on the sentences stating the time change themselves to the fact that working memory clearance causes additional processing when the stated change is large. This latter

result is very similar to the finding by Mandler and Goodman (1982), that the first sentence of a new story constituent tends to have a relatively long reading time; Mandler and Goodman suggested that longer reading times for topic change sentences may reflect the formation of a macro-proposition summarising the previous content. It should also be pointed out again that Kintsch (1977) suggested that explicit time changes were one of the cues readers used to identify structural boundaries.

Mandler and Goodman (1982) make an interesting observation concerning the effects of story category boundaries on pronominal reference. When writing their experimental materials, Mandler and Goodman found that they unintentionally used more full nouns in the first sentences of constituent nodes, and more pronouns in the second sentences of constituent nodes. This difference was attributed to the slight topic shift caused by transitions from one story unit to another. Evidence that pronominal reference across text units is unusual is also cited by Webber (1980), who briefly refers to a study which analysed the thematic structure of newspaper articles; this study found no instances of pronominal reference across thematic boundaries. However, Webber (1980) does not explain how thematic boundaries were determined in the texts studied, and clearly the structure of newspaper articles will be distinct from that of stories. Pronominal reference across story episode units is at least possible, since Mandler and Goodman (1982) were eventually able to equalise the number of pronominal references in the two sentences of the units which they studied. It is possible that pronominal reference may be still more uncommon across boundaries between complete episodes than across boundaries between episode components. In either case, presumably the presence of a structural boundary between pronoun and antecedent would lead to increased

difficulty of pronoun interpretation; no specific test of this appears to have been reported in the literature.

In addition to time changes, Sanford and Garrod (1981) speculate that explicitly stated spatial changes may also lead to loss of foregrounded status, if the change exceeds the normal limits of the current scenario. As with the effect of time changes, any such effect would not be due to overt statements that the scene had been changed, but to an inference based on stereotyped knowledge of scene limitations. It is not clear, however, whether scene changes produce a single discontinuity or may produce a cumulative effect as more and more changes occur; unlike the case of the distance effect, where this question has at least been studied, no experiments have been published comparing single scene changes with the effects of two or more scene changes. It is evident even from this one experiment, however, that availability of antecedent entities is under active strategic control. If loss of items from working memory were a simple function of capacity and duration limitations, the size of a stated time change would be immaterial; the fact that large stated changes produce a fall in antecedent availability relative to small stated changes is a clear indication that working memory contents are selected on the basis of cues in the text.

The effects of topic changes on ease of anaphora interpretation were studied by Lesgold, Roth and Curtis (1979). This paper is a central one in the study of foregrounding, and as the manipulation which it employed was also used in the present programme of research it will be described in some detail. Possibly the most important feature of the paper is that Lesgold et al. made a test of interactions between distance and topical changes, as well as investigating topical effects

themselves. Passages were written in which an anaphoric target sentence could appear either immediately after its antecedent or be separated from the antecedent by two or four intervening sentences. In the foregrounded conditions, the intervening sentences all dealt with topics closely related to the content of the antecedent sentence, as well as to the target sentence itself; in the backgrounded conditions, the intervening sentences dealt with topics unconnected to the specific information of the antecedent sentence. Judging from the example passage which Lesgold et al. present, these new topics did not move the text to different physical or temporal scenes, but simply emphasised features of the shared setting of the passage which were unconnected with the antecedent itself. This manipulation is therefore distinct from the scene change effects reported by Sanford and Garrod (1981). When there were two intervening backgrounding sentences between target and antecedent, these always dealt with a single new topic; when there were four backgrounding sentences, however, on half the occasions these dealt with a single topic while on the other occasions the first two sentences dealt with one topic and the last two sentences with a second topic. This manipulation therefore varied the number of topic changes between pronoun and antecedent, as well as the number of intervening sentences. The target anaphors were definite noun phrases rather than pronouns, although Lesgold et al. at times imply that more complex units of information than single entities may be subject to foregrounding.

Passages structured in the manner described above were presented to subjects on a computer screen. Rather than presenting entire passages one sentence at a time, Lesgold et al. chose to present all the passage content except the target sentence as a single display; the targets were then presented separately when subjects pressed a key. This

procedure, which was similar to that used by Clark and Sengul (1979), seems likely to have given undue emphasis to the target sentences, especially as these were always the last sentence in each passage, and it may therefore be questioned whether this affected subjects' reading strategies; nonetheless, several important results were found. Firstly, target reading times were significantly faster when the intervening sentences dealt with topics connected to the antecedent than when they dealt with new topics. In neither case did the intervening sentences mention the antecedent explicitly, yet in the foregrounded conditions anaphor interpretation was clearly easier. This result suggests that topical continuity acted to maintain the availability of the antecedent. A second interesting result showed that a single topic change was sufficient to reduce availability, with further changes producing no additional effects; when there were four intervening backgrounding sentences, target sentences were read equally slowly whether these dealt with one new topic or with two. Taken together, these results suggest once again that antecedent availability is under active control during comprehension, and indicate that loss of foregrounded status produces a uniformly low level of availability which is not readily reduced further. Possibly the most interesting results, however, concerned the effect of distance. It was found that reading times were completely unaffected by distance if the intervening material dealt with topics related to the antecedent; reading times with two or four foregrounding sentences were equal to reading times when the target immediately followed the antecedent. In addition, the number of intervening sentences which were present proved irrelevant in the backgrounded condition; reading times with two backgrounding sentences were significantly longer than when target and antecedent were adjacent, but adding two more backgrounding sentences did not produce any further reading time increase. These latter results suggest that distance

itself has no direct effect on antecedent availability, and it can therefore be argued that those studies described above which obtained significant distance effects did so owing to confounded topic changes.

Lesgold et al. (1979) also investigated the regaining of foregrounded status after a series of backgrounding sentences. They presented subjects with passages in which intervening backgrounding material was followed by sentences which repeated some or all of the information from the antecedent sentence; the anaphoric targets appeared after the repetition sentences. It was found that both complete repetition of the antecedent information, and repetition solely of the crucial antecedent noun, produced target reading times equal to conditions where target and antecedent were immediately adjacent; this result indicates that antecedents which have once been backgrounded can nonetheless be explicitly reintroduced into a text and acquire undiminished foregrounded status. More interesting, however, were the effects of repeating information which was contextually related to the crucial antecedent, but which did not mention it explicitly. In one experiment, adding a single such context repetition sentence after a series of backgrounding sentences was found to produce slight but non-significant facilitation of reading times; a further experiment found that adding two context repetition sentences instead of one did produce reading times equal to the adjacent condition. Thus, antecedent availability was increased simply by mentioning concepts that had been connected to it earlier in the text. This manipulation was very similar to the topical continuity manipulation described earlier; evidently, such continuity can preserve an item's foregrounded status, while later reversions to that topic lead to the item regaining such status.

Lesgold et al. (1979) relate these results to the short-term buffer, or working memory, storage of foregrounded information. When an anaphor immediately follows its antecedent, that item is directly accessible in working memory and hence assignment can proceed rapidly; the foregrounding effect of topical continuity shows that items can be maintained in such an activated state if cues in the text indicate their continued relevance. Topic changes lead to antecedents being cleared from working memory, and hence comprehension of subsequent anaphors incurs the added processing difficulty of reinstating the antecedent from the long-term text representation; context repetition sentences were seen as producing antecedent reinstatement before the targets were read, and so were equivalent to the other foregrounded conditions. A second explanation for the results was also offered, but discussion of this will be deferred until Chapter IV. For the moment, the important conclusion is that antecedent availability varies in ways which are responsive to text content and structure, and does not appear to be dependent on simple textual distance. However, it must be pointed out that the experiments by Lesgold et al. all made use of definite noun phrase anaphora; whether similar effects would be found in the case of pronoun interpretation has not yet been demonstrated. Furthermore, the emphasis which their procedure placed on the final target sentences may have led to unusual reading strategies which were not a true reflection of anaphor comprehension processes. Replication and extension of these effects would therefore be desirable.

There is thus considerable evidence that ease of anaphor interpretation is dependent on the foregrounded status of antecedents. Foregrounded antecedents are those which appear to be topical or relevant to the current text; such relevance can be lost if subsequent text implies a change of scene, as described by Sanford and Garrod

(1981), or if it simply emphasises other topics within the current setting, as shown by Lesgold et al. (1979). However, the experiments described so far have only been concerned with foregrounded status as it is affected by the text following the first mention of an entity; as noted above, there is also evidence that foregrounding can be affected by the manner in which an item is first mentioned. Furthermore, there are influences on antecedent status which appear to assign globally topical status to certain entities, such that they become less sensitive to topic changes occurring in the subsequent text.

One result illustrating effects of global topicality has already been mentioned. In the experiment on effects of distance described by Sanford and Garrod (1981), and discussed in the previous section, it was found that the distance effect on anaphor interpretation had different properties for antecedents in subject and object position. While anaphoric references to antecedents originally mentioned in sentence subject position did become harder to process as distance increased, target reading times were always faster for pronominal references than for definite noun phrases; with antecedents which were first mentioned in sentence object position, however, at the greater distance reading times were faster for anaphoric definite noun phrases than for pronouns. Sanford and Garrod explain this result by arguing that sentence subjects are taken as part of the main topic of discourse, and hence are retained in working memory over longer distances than sentence objects; references to such foregrounded entities by definite noun phrases are said to be inappropriate, and hence are read more slowly, while the reverse is true for backgrounded entities. Hence, topical status as signalled by the manner of an antecedent's first mention could override the topical changes occurring through a set of intervening sentences. The topical status of sentence subjects was also said to be reflected in

the fact that targets referring to such antecedents were always read faster than targets referring to sentence objects.

An effect of global topicality was also demonstrated in the experiment by Anderson on stated time changes (Sanford and Garrod, 1981). The passages used in that experiment in fact contained two crucial entities; one of these was the scenario-dependent entity which was found to become more difficult to refer to after a large stated time change, while the other was alleged to be the centrally topical character of the passage. Sanford and Garrod do not state how such global topicality was determined, but the example passages given suggest that these entities were always the first entity mentioned, and were then mentioned explicitly several times. The passages were sometimes presented with targets referring to the scenario-dependent entity, as described earlier, and sometimes with a target referring to the globally topical character. Two important results were found. Firstly, reading times for references to globally topical characters were always faster than for references to scenario-based characters; it was argued on this basis that the central topic had a dominant position in working memory. Secondly, and most importantly, it was found that references to central characters were unaffected by the size of the stated time change; reading times were equally fast whether the change was within or beyond the normal scenario limits. This result was seen as demonstrating that the time change effects were selective, and led to reduced availability of entities that were specifically tied to the current scenario, but not of entities which were more globally involved in the text. It is clearly plausible that implied scene changes should not be taken to indicate a fall in the relevance of entities that are likely to move between scenes.

These two results described by Sanford and Garrod are important not only because they show the selective effects of topic changes, but also because they suggest that variations in antecedent availability may be continuous rather than discontinuous. In the experiment on distance, subject antecedents were always more easily referred to than object antecedents, but nonetheless became more difficult to refer to with distance; while in the time change experiment, scenario-dependent characters were already more difficult to refer to than global characters when the time change was small, but became even less available when the time change was large. It was not the case in either experiment that two entities were equally available at one stage, but that one became less accessible owing to a topic change; rather, one entity appeared to be always more accessible than another, with topic change effects overlaid on this difference. Such a pattern of results is difficult to account for on the view that foregrounding is an all-or-none distinction, as is implied by results such as those of Clark and Sengul (1979) or Lesgold et al. (1979). Sanford and Garrod in fact suggest that several entities may be present in working memory at any one time, but that one of these predominates owing to its global topicality. However, the presented results may be explained more simply when it is noted that the target sentences referring to the two different entities were quite distinct, in both the distance and time change experiments. That is, a different set of sentences was used to refer to subject antecedents than was used to refer to object antecedents, and a different set of sentences was used to refer to central entities than to scenario-dependent entities. Thus, apparently greater ease of reference to one entity rather than another may have been due simply to differences in the inherent readability of the two sentence sets.

This criticism does not apply to the study by Cirilo (1981), which also found an apparent effect of global topicality. It has already been described how Cirilo varied the distance between a constant target sentence and its antecedent; however, Cirilo also varied the global relevance of that antecedent. This was done by analysing the experimental texts according to the story grammar presented by Thorndyke (1977); globally relevant antecedents were those which had a relatively high position in the hierarchical text representation, while less globally relevant antecedents were relatively low in the hierarchy. It was predicted that high-level antecedents would be maintained in working memory for longer than low-level entities, owing to their greater importance in the text; thus, high-level antecedents should have been less sensitive to the distance effect. The results did not support this view. Although target sentence reading times were found to be faster when the anaphoric noun phrase referred to a high-level entity rather than to a low-level entity, the distance effect appeared equally strongly in both cases. This outcome is clearly analogous to the finding that subject antecedents are more easily referred to than objects, but nonetheless become less available with distance.

Sanford and Garod (1981) explain such effects by arguing that items represented in working memory may be allocated different amounts of storage space, and that they may therefore differ in accessibility. Cirilo (1981), however, offers a quite different explanation. Broadly, Cirilo argues that effects of global relevance are not due to differential accessibility in working memory, but to the nature of the search process for antecedents. If the search process begins from higher levels of the text representation, high-level entities will necessarily be most easily located. This possibility will be left on one side for the time being, but will be returned to in much more detail

in Chapter IV.

The preceding discussion of foregrounding can be summarised as follows. A number of studies have shown that certain ways of mentioning an entity lead to that entity either being easier to refer to again (Yekovich et al., 1979), or being the preferred antecedent for ambiguous pronouns (Carpenter and Just, 1977a; Ehrlich, 1980). These effects, and the ideas of Chafe (1972), could be unified in the theory that topical entities were intrinsically easier to refer to again; this ease of reference has been attributed in psychological theory to antecedent presence in working memory. Various results have shown reductions in antecedent availability to be caused by distance, stated time changes, and introductions of new topics; while globally topical entities may be generally more accessible than less globally topical entities. The broad explanation most frequently offered for these effects on availability is to claim that pronouns can be directly matched to foregrounded entities in working memory, but can only be assigned to backgrounded entities if they are first reinstated from the long-term text representation.

This view of pronoun interpretation, which is best exemplified by the detailed presentation given by Sanford and Garrod (1981), argues that successful assignment can be achieved in two distinct ways. Since writers normally produce texts that are cooperatively structured, such that topical continuity is preserved and pronouns are used to refer only to currently foregrounded entities, it is argued that pronoun interpretation processes can initially be extremely simple in operation. When a pronoun is encountered, readers are seen as making an attempt to assign it directly to some entity present in working memory; this first method of assignment is usually termed immediate matching or direct

mapping. Matches made in this way are immediate in two respects. Firstly, direct availability of an antecedent in working memory allows the connection to be made easily; and secondly, this process is said to be attempted as soon as the pronoun is encountered. Sanford and Garrod (1981) make this latter assumption explicit, and regard a pronoun as an instruction to the reader to retrieve an antecedent from working memory solely on the basis of the content of the pronoun itself. Pronoun interpretation by means of such primary processing will normally be successful, since antecedents will usually be present in working memory. However, if a text is badly structured, or if for some other reason the intended antecedent to a pronoun is not currently activated when the pronoun is encountered, then the immediate matching strategy will fail. In such cases, pronoun assignment will be attempted by the second method. This process is more complex than the first, and involves searching for the antecedent in the long-term text representation. These retrieval searches are seen as making use of much more information than the content of the pronoun itself, by drawing on other disambiguating content and making use of pragmatic inferences.

It is evident that antecedent presence in working memory is only an explanation of the foregrounding phenomena described above if pronoun interpretation is conceived in this highly restricted way. Longer reading times in the various backgrounded conditions are explicable only on the view that immediate matching is always attempted when a pronoun is encountered; such matching succeeds in the foregrounded conditions, but in the backgrounded case a second processing stage is executed after the first has failed. It is this additivity of processing stages which forms the basis of the predicted longer reading times in backgrounded conditions. Although retrieval processes are normally assumed to be intrinsically slower in operation than direct mappings, this difference

cannot be guaranteed on intuitive grounds alone; it is therefore necessary to assume that retrieval processes are not commenced until direct mapping has failed. If this were not the case, retrieval processes in backgrounded conditions could in principle operate as rapidly as direct mapping processes in foregrounded conditions.

As Sanford and Garrod (1981) make clear, this distinction between primary and secondary processing forces a further distinction between the types of influence which may affect pronoun interpretation at each stage. During attempted direct matches, the assignment process is assumed to make use only of the lexical content of the pronoun and of deterministic restrictions based on syntactic structure; the weight of the processing explanation for correct assignments in the absence of such determining biases therefore rests on the appropriate prior selection of working memory contents. Semantic disambiguation and pragmatic inferences based on event relations and world knowledge are called into play only during secondary retrieval processing, should this stage become necessary. In cases where the intended antecedent to a pronoun is correctly foregrounded, such complex sources of information are seen as playing no role in assignment.

The working memory explanation of foregrounding therefore makes very strong assertions concerning the nature of assignment processes. Various features of text structure and content clearly influence the accessibility of items to pronominal reference, but these influences are seen as affecting working memory presence rather than as operating on the assignment processes themselves. Assignment processes are seen as occurring in an invariant sequence of stages, irrespective of the pronoun's context of occurrence; the initial direct mapping stage in particular is identical in all cases of successful primary processing

assignment. In addition, initial assignment attempts are seen as simple, restricted matching processes; the complex influences on pronoun assignment play their part only in the processing decisions made in advance of encountering the pronoun itself.

The first set of experiments in this research was carried out within the framework of the working memory theory, and was intended to clarify such issues as the nature and sources of the distance effect, the consequences of topical continuity for pronoun interpretation, and the relative informativeness of the reading time and pronoun question measurement techniques. Before these experiments can be described, however, one further aspect of this thesis must be introduced.

In addition to its intrinsic interest as a feature of human language use, one important reason for studying reading is its great practical relevance. Reading is ubiquitous in everyday life and is a vital medium of information exchange. Since reading deficiencies can therefore be a handicap in several ways, it is important that the reasons why certain individuals fail to achieve acceptable reading standards should be uncovered; understanding the sources of low levels of reading skill should assist in remedying those deficiencies. In addition, even among readers of normal ability the discovery of ways in which reading efficiency could be increased would be of great practical benefit. The study of individual differences in reading ability also interacts closely with the study of reading as a topic in its own right. In order to explore possible sources of individual differences, it is necessary to have an overall theory of reading which indicates the types of problem which may occur and their most probable processing explanation; while conversely, any valid theory of reading must be able to accommodate known facts about individual differences. The study of individual differences thus offers an additional method of testing psychological theories of reading.

A great many suggestions have been made in the literature concerning possible causes of low reading ability levels, and reviews may be found in Carr (1981) and Mitchell (1982). The present review, however, will be restricted to suggested differences between good and poor readers which are relevant to pronoun interpretation. These suggestions revolve around alleged differences in working memory capacity and in sensitivity to text structure. The theory of pronoun

interpretation described above suggests that working memory contents must be controlled in such a way as to ensure that probable antecedents are present there, while the cues used to guide such working memory selection are derived from aspects of text structure. If poor readers were found to be deficient in either of these skills, this should produce consequences for their pronoun interpretation processes; while at the same time, pronoun interpretation could be used as a research tool to investigate possible skill differences in the factors which influence it.

As was noted earlier, Kintsch and van Dijk (1978) suggested that poor readers may suffer from a smaller working memory capacity than good readers, and hence would be able to carry over fewer propositions from one processing cycle to the next. The consequences of buffer size differences for ease of reading were demonstrated by Kintsch and Vipond (1980), who analysed various texts on different assumptions concerning buffer capacity. As buffer size changes, so too does the number of occasions on which each new input cycle fails to find a match among the currently active propositions, with corresponding changes in the number of time-consuming reinstatement searches which would be required for comprehension. The interesting finding was that the theoretically predicted effects of buffer size changes varied across passages. While one passage which required a large number of reinstatements with a small assumed buffer capacity was found to require no reinstatements at all when capacity was increased, a second passage required equal numbers of reinstatement searches whether buffer capacity was assumed to be small or large. As Kintsch and Vipond (1980) point out, such results indicate that individual differences in working memory capacity need not be associated with uniform effects on reading ability, but would have specific consequences for the readability of different texts.

Kintsch and Vipond (1980) in fact presented no data comparing the working memory capacity of skilled and less-skilled readers, but their theoretical analysis shows that such differences would be a possible cause of individual variations in ability. Empirical evidence that poor readers do suffer from smaller working memory capacity than good readers was provided by Perfetti and Goldman (1976). This study used auditory presentation of short stories, which were interrupted on occasion by repetition of a probe word from earlier in the passage; the subjects' task was to say the word which had immediately followed the probe word in the text. It was argued that performance on this probe discourse task would be superior if the required words were present in working memory when the probe was presented. Subjects were third and fifth grade school children, divided into skilled and less-skilled reader groups on the basis of scores on a standardised reading test. The two skill groups were matched on IQ. It was found that the less-skilled readers recalled significantly fewer target words than skilled readers. The intervention of a sentence boundary between the occurrence of the target word in the text and the presentation of the probe word was found to reduce accuracy, thus confirming the results of Jarvella (1971) indicating rapid loss of verbatim content once a sentence boundary was crossed; the two reader groups were equally sensitive to the boundary effect.

The findings of Perfetti and Goldman (1976) are relevant to working memory differences between good and poor readers even though auditory presentation was used; as noted earlier, it is generally assumed that reading and listening comprehension share common features at a sufficiently deep level of processing. However, Goldman, Hogaboam, Bell and Perfetti (1980) have in fact extended the probe discourse results to reading. The child subjects in that study read passages printed in

small booklets with only 20-30 words per page; turning a page occasionally revealed a single probe word, with subjects again being required to say the word which had immediately followed it in the preceding text. It was found once again that less-skilled readers had poorer recall of target words than skilled readers. Goldman et al. (1980) did however find some interesting results concerning the effects of sentence boundaries. In their first experiment, it was discovered that target recall was reduced if the test occurred after a sentence boundary had been crossed, but only if a meaningful unit of the succeeding sentence was already being processed; recall was unaffected by a sentence boundary if only the first word of the next sentence was read before the probe word was presented. In their second experiment, Goldman et al. increased the length of the text segments intervening between the target word and presentation of the probe, in the expectation that with extremely long segments readers would be unable to maintain the entire current sentence in working memory. In accordance with this prediction, it was found that recall performance with long intervening segments was reduced compared to that with short segments even in the absence of a sentence boundary. It was also found that while skilled nine year old readers showed a sentence boundary effect on long segments, twelve year old subjects did not; it was argued that the older readers had acquired strategies for controlling the rate of information input which did not rely on sentence boundaries, while the younger readers were continuing to attempt to use sentential units even when these were too large for this to be efficient.

There is therefore evidence from both reading and listening studies that less-skilled young readers show poorer verbatim retention during discourse comprehension. The sources of this deficit are discussed by Perfetti and Lesgold (1977, 1979). Perfetti and Lesgold argue that the

poor reader working memory deficit is specific to discourse processing. They suggest that, rather than having smaller absolute short-term memory capacity, less-skilled readers have only smaller effective capacity in verbal processing. The cause of this smaller effective capacity is said to be deficient word coding. There is evidence from word identification studies (Perfetti and Hogaboam, 1975; Perfetti, Finger and Hogaboam, 1978) that less-skilled readers are relatively slow at decoding individual words; in Perfetti and Lesgold (1977) it is argued that this could cause slower coding of information into and out of working memory. As Carr (1981) points out, however, this argument would in fact predict that poorer readers should show less rapid loss of information from working memory, rather than more rapid loss as Perfetti and Lesgold wish to prove. A more plausible argument is put forward in Perfetti and Lesgold (1979). Here, the interactive, shared-capacity view of reading is used to argue that inefficient word coding processes may occupy excessive central processing capacity and so reduce its availability for other purposes. If capacity is occupied with word recognition, less will be available for storage of coded words and earlier information will therefore be rapidly lost as later input is processed. Evidence that poorer readers do not suffer short-term deficits in non-verbal tasks was presented by Perfetti and Goldman (1976), who found no skill differences in performance on a probe digit task; analogously to the probe discourse task, this involved listening to a series of random digits and saying what number had followed a subsequently presented probe. However, in order to support their contention that working memory deficits are discourse specific, Perfetti and Lesgold (1977) are forced to explain away the large body of results indicating that poor readers do show smaller short-term memory span in more traditional measurement procedures; these explanations are not all wholly convincing. For example, Perfetti and Lesgold (1977) state that

conventional digit span measures have produced inconsistent results, and also argue that WISC digit span shows no relation to reading skill if IQ is controlled. Both these claims are contradicted by Nelson and Warrington (1980), who refer to the frequently replicated finding that poor readers show smaller digit spans than good readers and who themselves demonstrated this effect on less-skilled readers of above average IQ.

Whether the poor reader working memory deficit was general or specific to verbal comprehension was one of the questions studied by Daneman and Carpenter (1980). It has already been noted that Daneman and Carpenter found accuracy on questions interrogating the referents of pronouns to fall with increasing distance; the validity of this linear distance effect was questioned on the grounds of the unsatisfactoriness of the measurement technique used. However, Daneman and Carpenter also related pronoun question accuracy to reader skill and to measured working memory capacity. Subjects were given two tests of memory span. One of these involved listening to sets of unrelated words, and attempting to repeat them back in order of presentation; this was therefore a fairly conventional span test. The other test, however, was designed specifically to tax both the processing and storage components of reading. It was argued, in a similar way to Perfetti and Lesgold (1977), that poor readers may be relatively inefficient at reading processes and that this may reduce the capacity available for storage; such capacity reductions would not appear on a conventional span test with a minimal processing component. The reading span test which Daneman and Carpenter devised required subjects to read sets of sentences, presented one sentence at a time, out loud; at the end of each set, subjects attempted to repeat back the sentence-final words in presentation order. Span level was defined as the largest set size on

which a subject achieved perfect recall on at least two out of three different trials. Clearly, as subjects read the later sentences in a set they would also be attempting to maintain the final words from earlier sentences; inefficient reading processes would reduce storage space available for such retention, and so should produce lower span levels.

It was found that single-word span and sentence reading span showed different relations to reading ability and pronoun question accuracy. Firstly, word span showed no significant correlation with scores on a standardised verbal aptitude test; reading span, however, showed a significant positive correlation with this overall measure of reading ability. Secondly, while reading span showed a strong positive correlation with pronoun question accuracy, word span showed a much smaller and non-significant correlation. When an analysis of variance on pronoun accuracy was carried out with subjects divided into groups on the basis of their reading span levels, this measure was found to interact with distance; low-span subjects showed steep falls in accuracy even over relatively short distances, medium-span subjects had reduced accuracy only at the longest distances, while the highest-span subjects had perfect accuracy scores at all distances. Thus, those subjects with the smallest measured working memory capacities were most sensitive to the distance effect on pronoun interpretation. In a second experiment, span levels were calculated on both read and heard sentence sets, and compared with pronoun question accuracy on both read and heard passages. Reading and listening span were well correlated, while span level interacted with the effects of pronoun-antecedent distance on both reading and listening comprehension. These results were seen as indicating that any processing deficits of poor readers must be at a level common to both reading and listening; this would argue against

the view of Perfetti and Lesgold (1977, 1979), that poor readers are primarily deficient at visual word decoding.

In both their experiments, Daneman and Carpenter (1980) found that low-span subjects were less accurate on factual comprehension questions as well as on pronoun referent questions. It is unclear, however, whether the results as a whole are best used in support of the working memory theory of reading ability or of the working memory theory of pronoun interpretation. The fact that span level predicted performance on pronoun questions supports the view that ease of pronoun assignment is dependent on antecedent presence in working memory; yet Daneman and Carpenter evidently assumed that the working memory model of pronoun assignment was correct, and were more concerned with validating the working memory theory of individual differences by using pronoun interpretation as a way of testing this. If therefore pronoun interpretation is not dependent on working memory presence in the way that Daneman and Carpenter assume, or if as was suggested earlier the question technique does not reflect assignment processes alone, the correspondence between span level and pronoun performance must be given some alternative explanation.

Daneman and Carpenter's study was in any case founded on the assumption that antecedent loss from working memory should be at least in part a linear function of surface distance; as already noted, this assumption cannot yet be considered proved. There is clear evidence that items can be reduced in availability to pronominal reference in response to structural cues in the text, and it is possible that obtained distance effects are due to confounding with such cues. If this is the case, differences between good and poor readers in their pronoun interpretation processes would not be predicted solely on the

basis of working memory capacity differences, but also on the basis of differences in sensitivity to the textual cues used to guide selection of working memory contents. A poor reader capacity deficit may produce poorer overall pronoun interpretation performance, but would not be sufficient alone to produce differential sensitivity to effects of topical continuity, scene changes, and the like. If the distance effect is due to active clearing of working memory contents in response to structural or semantic boundaries, poor readers would be predicted to show greater effects of distance only if they were more sensitive to such discourse cues than good readers. This suggestion is certainly counter-intuitive, and there is in fact research evidence showing poor readers to be less sensitive to the relevant discourse cues than good readers.

Some of this evidence was presented by Daneman and Carpenter (1980) themselves. In their second experiment, subjects were asked to provide titles for each passage that captured the main theme; these titles were then rated according to how well they did in fact express the main idea. It was found that low-span subjects produced titles which often failed to capture the passage theme; mean ratings were lower for poor readers than for good readers. Other evidence that poor readers are less able to identify the main topic of a passage was presented by Eamon (1978-79). In that study, student subjects divided into skilled and less-skilled reader groups on the basis of a standardised reading test read short paragraphs which dealt with two distinct concepts; one of these was the thematic concept, which was repeatedly referred to in many different propositions, while the other was a non-thematic concept only incidentally referred to. In Eamon's first experiment, subjects read the passages and then rated statements from them according to how important that information had been to the passage. It was found that

while skilled readers gave higher importance ratings to statements about the thematic concept than to statements about the non-thematic concept, less-skilled readers showed no differences in ratings of the two statement types. The second experiment used a recall measure, and found that while skilled readers had superior recall for thematic information than for non-thematic, less-skilled readers had equal and low recall for both types of material.

Evidence that skilled and less-skilled readers differ in their sensitivity to text structure was presented by Meyer, Brandt and Bluth (1980). Subjects of approximately fifteen years old read non-narrative texts which were well organised in terms of standard rhetorical structures; for example, a passage dealing with oil spills from supertankers was cast as a problem statement plus solution. After reading the texts, subjects gave written recall. It was hypothesised that skilled readers, as identified by a standardised test, would tend to make more use of text structure than less-skilled readers. As a result, while the recall protocols of skilled readers should be structured in a similar way to the original texts, the protocols of less-skilled readers should consist largely of list-like sets of unrelated descriptions. Analysis of the recall protocols showed that the majority of skilled readers did in fact make use of inherent text structure, while the majority of poor readers did not. The advantage of using text structure was evident in the fact that those subjects who did use the structure strategy had superior recall.

The above results therefore suggest that poor readers are less able to identify the central topic of a text than good readers, and also make less use of inherent text structure. Perfetti and Lesgold (1977) acknowledged that such differences in sensitivity to text-level

variables could possibly be a source of variations in reader skill, but argued that the then existing evidence suggested that no such differences existed; these more recent results contradict that assertion. In addition, some of the experiments which Perfetti and Lesgold cite in support of their assertion are in fact plainly opposed to it. For example, an unpublished study by Berger tested children's recall of passages and their accuracy on literal questions about the content; it was argued that recall performance should be affected by sensitivity to passage organisation, while literal question accuracy would depend only on low-level comprehension. According to Perfetti and Lesgold, differences between skill and less-skilled reader groups were found to be substantial on both types of test. Even though this result therefore indicates that skilled readers were making more use of text structure, as well as being superior on literal comprehension, Perfetti and Lesgold cite the study in support of the exactly opposite contention.

The differences in topic identification and structure use between good and poor readers have clear consequences for pronoun interpretation. Text items are said to be foregrounded with respect to pronominal reference because of their topical status, while the crossing of discourse structure boundaries may be one way in which foregrounding is lost; however, poor readers would presumably be less likely to correctly identify the currently topical concept, and less likely to change the identified topic when a text boundary was encountered. Thus, while good readers efficiently select and adjust their working memory contents in response to topical cues in the text, poor readers may either not make use of a topic identification strategy at all or may operate it very inefficiently. While poor readers may therefore experience generally greater difficulty at pronoun interpretation than

good readers, it can be predicted that poor reader sensitivity to foregrounding manipulations should be less than that of good readers. Good readers should show the expected sharp fall in ease of pronoun assignment after a topic change, while poor readers should not.

This prediction is clearly at variance with the results of Daneman and Carpenter (1980), who found that less-skilled readers were more sensitive to the effects of distance on pronoun interpretation than skilled readers. There are several ways in which this disagreement can be explained. Firstly, it must be pointed out that Daneman and Carpenter in fact divided their subjects into groups on the basis of their span level rather than their reading ability; while the two measures had a significant correlation of approximately +0.6, the span division is clearly not a direct test of skill differences. Secondly, Daneman and Carpenter's reading ability measure consisted simply of asking subjects their scores on the Verbal Scholastic Aptitude Test; it is not stated how recently these tests had been taken, who had administered them, or how the dependability of subjects' answers was assessed. Thirdly, as noted earlier the pronoun question technique may reflect inferences made when answering, as well as ease of pronoun assignment; as will be discussed in connection with the present Experiment 2, poor readers may be less efficient at drawing inferences.

The difficulty of reconciling the results of Daneman and Carpenter (1980) with the prediction that less-skilled readers should be less sensitive to foregrounding may also be overcome by considering the source of the distance effect. As noted earlier, Daneman and Carpenter predicted a linearly increasing effect of distance on pronoun interpretation, while results such as those of Lesgold et al. (1979)

suggest that distance effects may be consequences of confounded topical boundaries. It is on the latter view that poor readers, being less sensitive to effects of text structure, would be predicted to be less sensitive to the distance effect than good readers. Only the former view, attributing the distance effect to general processes of working memory loss based on decay and displacement, would predict greater sensitivity of poor readers. However, the interaction between distance and topic changes has not yet been studied specifically with pronouns; it is possible that pronominal anaphora, unlike the definite noun phrase anaphora studied by Lesgold et al. (1979), would be found to show a residual distance effect in addition to effects of topical changes. If this were the case, greater sensitivity of poor readers to the distance effect would in fact be predictable on the basis of working memory capacity differences, as argued by Daneman and Carpenter (1980).

The question of skill differences in pronoun interpretation is therefore confused both by uncertainty regarding the nature and sources of the distance effect, and by uncertainty regarding the validity of the pronoun question technique. The first set of experiments in this research included investigations of differences between good and poor readers, in an attempt to clarify these problems.

CHAPTER II

Distance and Foregrounding Effects on Pronoun Interpretation

- 1 Experiment 1
- 2 Experiment 2
- 3 Experiment 3
- 4 General Discussion

Introduction

The preceding overall introduction and literature review described the principal issues concerning pronoun interpretation and sources of individual differences in reading ability; Experiment 1 was designed to bring these two areas together, in an attempt to test questions of interest to both.

One of the topics to be addressed by this experiment was the effect of pronoun-antecedent distance on pronoun sentence reading times. Clark and Sengul (1979) found that reading times increased when the sentences containing pronoun and antecedent were separated by one other sentence; the present experiment also attempted to demonstrate this effect, while making some improvements in procedure which were intended to increase the validity of the results. One problem with Clark and Sengul's study, for example, was their failure to give subjects a comprehension test on the material they read, thus leaving it unclear whether they fully attended to the content. This problem was easily corrected in the present experiment, by giving subjects comprehension questions on each passage. A more important problem with Clark and Sengul's study was the fact that the target pronouns, although intended to refer to one specific entity, may in fact have been assigned to other referents by their subjects; this potential ambiguity may be clearly seen in their example passage, where the target sentence 'It appeared to be an antique' could refer to either a brass lamp or an upholstered chair. This possibility is important because of Clark and Sengul's claim that reading times were no longer when two sentences separated pronoun and

antecedent than when they were separated by only one sentence; if in the former case pronouns were sometimes being assigned to more recent referents than those intended, any effect of increasing distance would have been reduced. In the present experiment, materials were constructed such that the target pronouns could only be assigned to the intended antecedent. Thus, although it was predicted that pronoun sentence reading times would increase with pronoun-antecedent distance, no prediction was made as to whether this increase would be continuous or would reach an eventual asymptote; discovering whether or not the distance effect was linear and increasing would allow firmer conclusions to be drawn concerning the sources of that effect.

The second question of interest in this experiment concerned the possible interaction of the distance effect with reader skill. Daneman and Carpenter (1980) discussed the reasons why poor readers should be more sensitive to the effects of distance on ease of pronoun interpretation, and the results using their pronoun question technique offered support for this view. The present experiment made a further test of this suggested differential sensitivity to the distance effect, using the reading time technique to obtain a more direct measure of local processing difficulty. In this way the problems involved with the pronoun question technique were avoided. In line with the arguments of Daneman and Carpenter (1980), it was predicted that poor readers would show a greater sensitivity to the distance effect than good readers.

A common technique in studies of reading ability is to divide a single subject sample into two groups at the median of some reading test administered to only those subjects, and to label the high-scoring subjects good readers and the low-scoring subjects poor readers. This approach has advantages of simplicity and practicality, but does not

guarantee a wide difference between the two skill levels; subjects close to the median may in fact be very similar in ability. To avoid this danger, in the present experiment a large sample of subjects was pretested using a standardised reading test, and the best and the worst readers from this group were used in the main experiment. In this way, a clear difference between the skilled and less-skilled reader groups was made more likely.

One other feature of the experiment must be mentioned. The most usual technique in reading time studies is to present each display separately, with only one display being visible at a time; when concerned with text integration, this procedure clearly forces a reliance on the memory functions which are the topic of interest here. However, it is also possible that this method removes one cause of reading time increases. Carpenter and Just (1977a) showed that regressive eye fixations occurred on roughly half the occasions a pronoun was encountered; if these regressions occur only when an antecedent is not directly available in working memory, and act as an aid to retrieval, they would be predicted to be more likely at greater pronoun-antecedent distances. This would lead to a further elevation of reading times, in addition to more prolonged gazes on the target sentences themselves. The inability to make such regressions with sentence by sentence presentation may cause subjects who realise they have no antecedent directly available to proceed immediately to the next display, since they cannot pursue normal recovery procedures. This would serve to weaken any reading time effect. Since reading skill was an important factor in this experiment, and poor readers are known to make more regressions than good readers (evidence cited by Carpenter and Just, 1977a), it was decided that cumulative presentation of text sentences would be a more sensitive technique than clearing past

sentences before subsequent ones were displayed. Earlier sentences were thus left in position as subjects read later ones. It may be noted that this technique is a closer approximation to the conditions of normal reading, while still allowing measures of immediate processing duration.

In summary, it was predicted in this experiment that pronoun sentence reading times would increase with pronoun-antecedent distance, possibly to an eventual asymptote; and that skilled and less-skilled readers would differ in sensitivity to this effect.

Method

Subjects

Twelve skilled and twelve less-skilled readers were required for the main experiment, and to this end a much larger group was screened for reading ability.

The subjects in the screening test were 74 first-year students taking an introductory psychology course at Plymouth Polytechnic. They were given the comprehension section of the Nelson-Denny test (Form A) in a single group session following a lecture. They were told that a subgroup would be selected on the basis of their performance on the test, but were not told that good and poor readers were being chosen.

The Nelson-Denny test provides a score for both reading rate, measured in words per minute, and comprehension accuracy, measured in number of correct answers to multiple-choice questions. In the group of

74, the scores on these two measures had a product-moment correlation of +0.137, which was not significant. Despite this absence of correlation, however, it did prove possible to select groups of subjects who scored either high or low on both measures. Subjects were ranked in order, independently for each measure, and only those subjects with either high or low ranks on both were chosen. Relative to the 74-subject group, the twelve subjects selected as skilled readers were all in the top 39% of subjects on reading rate and the top 36% on accuracy, while the twelve subjects selected as less-skilled readers were all in the bottom 40% of subjects on reading rate and the bottom 43% on accuracy. The selected subjects were the only ones for whom both scores fell within the indicated cut-off ranges; these ranges were the result of the selection procedure, not deliberately chosen standards. In the selected group of 24 subjects, the two measures had a correlation coefficient of +0.781 ($df=22$, $p<0.005$), which confirms the effectiveness of the selection procedure used.

The selected subjects took part in the main experiment approximately two to three weeks after the screening session. Only these subjects were paid, receiving £2 for their participation.

Materials

Eighteen passages were written for the experiment. These were six sentences long, and described brief narrative episodes in familiar locations, such as a shop or a cinema. Each passage could be varied to take any of six possible forms, which corresponded to six experimental conditions. An example passage is presented in Table 1.1, with sentences labelled as described below. All eighteen passages are given

Table 1.1

Experiment 1: Example Passage

First	It was the final day of the race meeting
Filler 1	The racecourse was very busy
Filler 2	The stands were filling rapidly
Antecedent	A jockey fell at the second fence
Target	He/A man cried out in shock
Last	One of the following horses also fell
Question	Did a large number of people see the accident?

The individual passages were constructed as follows. Each passage contained a sentence which introduced a particular individual by means of an indefinite noun phrase, always in sentence-initial position. This sentence will be termed the antecedent sentence. Thus, in the example passage, the crucial antecedent entity is the jockey. A later sentence, which will be called the target sentence, could appear in either of two variant forms. In the first variant, the target sentence referred to the individual introduced in the antecedent sentence by means of an unambiguous sentence-initial pronoun. The pronoun 'he' was used 12 times, and 'she' was used 6 times; these pronouns were unambiguous in that the person referred to in the antecedent sentence was the only individual mentioned in the passage. In the example, if the target sentence was presented with 'he' in sentence-initial position, this would clearly refer to the jockey. In the second possible variant, in place of a pronoun the target sentence contained an indefinite noun phrase which introduced a second individual to the passage. The noun phrases used were 'a man' and 'a woman'; if the target sentence in the

first variant form included 'he', the second variant substituted 'a man', and similarly for 'she' and 'a woman'. The passages were written such that either of the two variants of the target sentence, which differed only in whether a pronoun or the corresponding noun phrase occurred in sentence-initial position, would be equally appropriate to the context. In the example passage, it can be seen that it would be equally plausible for either the fallen jockey or some other man watching the incident to cry out in shock. Whether the target sentence contained a pronoun or an indefinite noun phrase will be termed the Reference Type factor. It was expected that presence or absence of the antecedent in working memory would only affect ease of interpretation of the target sentence when it contained a pronoun referring to that antecedent; the inclusion of the variant containing an indefinite noun phrase was thus a form of control.

In addition to the antecedent and target sentences, each passage contained four other sentences. The first and last sentences in each passage were constant across all six forms, while the other two sentences, which will be termed filler sentences, were varied in position relative to the antecedent and target sentences. Three orderings of these sentences were used, in order to vary the number of sentences that intervened between the target and the antecedent. These orderings are given in Table 1.2. In these constructions, it can be seen that while the antecedent-target distance varies, the position of the target sentence in the passage remains the same. The passages were written such that any of these three arrangements would be equally possible, whether the target sentence contained an indefinite noun phrase or a pronoun. The number of sentences intervening between target and antecedent constituted the Distance factor, and as indicated in the table had levels of 0, 1, and 2.

Table 1.2

Experiment 1: Schematic Representation of Three Passage Orderings

	<u>PASSAGE ORDERING</u>		
	<u>1</u>	<u>2</u>	<u>3</u>
<u>SENTENCE</u>	First	First	First
<u>LABELS</u>	Filler 1	Filler 1	Antecedent
	Filler 2	Antecedent	Filler 1
	Antecedent	Filler 2	Filler 2
	Target	Target	Target
	Last	Last	Last
<u>PRONOUN</u>			
<u>ANTECEDENT</u>	<u>0</u>	<u>1</u>	<u>2</u>
<u>DISTANCE</u>			

In order to permit the various rearrangements described above, the content of the filler sentences was kept as neutral as possible. They described general features of the setting, such as the weather, or dealt with predictable features, such as the shelves of books in a library. They contained no references to specific new individuals, and nor did they refer to the person introduced in the antecedent sentence. In addition to this control over their content, the filler sentences were also standardised in length. Each filler sentence contained five words and eight syllables. The portion of the antecedent sentence following the initial noun phrase was also standardised in length, always containing five words and six syllables. These standardisations ensured that the distance manipulation would be comparable in all eighteen passages. Although the other sentences in the passages were of a similar length to the antecedent and fillers, they were not rigidly controlled.

A single yes-no comprehension question was written for each passage, and was used whichever form the passage happened to take on a particular presentation. Half the questions required the answer 'yes' and half the answer 'no'. These questions did not interrogate the referent of the target pronoun, not only because of the difficulties in interpreting such a measure but also because questions of that type might possibly have caused subjects to realise the importance of the pronouns in the experimental materials.

In summary, there were six experimental conditions, defined by the crossing of the Reference Type and Distance factors, and any passage could be varied to fit into any condition. Twelve random sequences of the passages were constructed. Each sequence contained all eighteen passages, three in each experimental condition. The decision on which form each passage took in each sequence was made randomly, with the constraints that there were three passages of each type in each sequence, that no passage was repeated in any one sequence, and that each form of each passage occurred in two different sequences. Within each sequence, the order of the passages was randomised. Each subject was presented with one sequence; each sequence was presented to both one skilled and one less-skilled reader.

In addition to the experimental passages, two practice passages were written. These were similar in length to the experimental passages, but were not constructed in the same rigid format. They never varied in form, and were presented to all subjects. The practice passages are contained in Appendix 1.

Procedure

Subjects were tested individually in a session lasting approximately thirty minutes, including instructions and debriefing.

Subjects were seated at a comfortable viewing distance in front of the visual-display unit of a DEC-MINC 11 mini-computer, which controlled the experiment. They were told that the experiment was concerned with reading comprehension, and that they were going to see a number of short passages displayed on the screen, followed by questions. They were led to believe that the questions were the main purpose of the experiment.

Before each passage, a warning message was displayed on the screen, telling subjects to press a red button, on the left of a separate response-key board, when they were ready to begin. When this button was pressed, the warning message was replaced by the first sentence of the passage, which appeared in the middle of the screen. The sentences were displayed in lower-case, with sentence-initial capitals and sentence-final full stops; the computer display was of excellent quality, with all required ascenders and descenders on each letter. The subsequent sentences of each passage appeared directly below the first, on successive lines with a straight left margin. Previous sentences in a passage remained on the screen as the later sentences were displayed. Subjects were instructed to press the red button as soon as they had read and understood each sentence, in order to bring up the next. The reading times for each sentence, defined as the interval between successive button presses, were recorded automatically by the computer.

When the red button was pressed after the last sentence of each passage, the entire text disappeared and was replaced by the question. To answer the question, subjects were told to press either of two keys on the right of the response board, labelled respectively 'YES' and 'NO'. If they were unsure in any particular case, they were instructed to pick whichever seemed most likely. When a response had been given, the question disappeared from the screen, and after two seconds the warning message for the next passage was displayed. Question answers and response latencies were recorded automatically by the computer.

Subjects were given the instructions orally. After reading the two practice passages in the presence of the experimenter, and showing that they understood the procedure, they were left alone to go through the experimental passages.

Results

Table 1.3 presents the mean target sentence reading times in milliseconds for each of the six experimental conditions, for both skilled and less-skilled readers. A three-way analysis of variance was carried out on these data, with Skill as a between-subjects factor and Distance and Reference Type as within-subjects factors. As there were three observations per condition for each subject, the analysis was conducted on the means of these three values.

Table 1.3

Experiment 1: Mean Target Sentence Reading Times (msec)

	<u>REFERENCE TYPE</u>						
	<u>PRONOUN</u>			<u>NOUN PHRASE</u>			
	<u>DISTANCE</u>			<u>DISTANCE</u>			
	<u>0</u>	<u>1</u>	<u>2</u>	<u>0</u>	<u>1</u>	<u>2</u>	
<u>SKILLED</u> <u>READERS</u>	2084	2485	2823	2455	2516	2192	2426
<u>LESS</u> <u>SKILLED</u> <u>READERS</u>	3152	2921	3058	3286	3188	2741	3058
	2618	2702	2940	2871	2852	2466	

The analysis, for which the ANOVA summary table is given in Appendix 14, showed that there were no significant main effects or interactions. It was found that skilled readers had a mean target sentence reading time of 2426msec, while less-skilled readers had a mean of 3058msec; this difference, however, although in the predicted direction, did not prove significant, with $F(1,22)=1.71$. For both the main effect of Distance and the main effect of Reference Type, F values were less than 1. The same was true for the Reference Type x Skill interaction, and for the three-way Distance x Reference Type x Skill interaction; while for the Distance x Skill interaction, $F(2,44)=1.07$. The only effect to approach significance was the Distance x Reference Type interaction; for this interaction, $F(2,44)=2.61$, $p<0.1$. From Table 1.3 it can be seen that while reading times for target sentences containing pronouns did tend to increase as target-antecedent distance

increased, the reading time for target sentences containing noun phrases fell by almost 400msec when there were two intervening sentences. The increase in reading time for pronoun sentences was a central prediction of the experiment, but the drop in reading time for noun phrase sentences was unexpected.

The overall analysis thus provided little support for any of the predictions made earlier. The only encouraging result was the Distance x Reference Type interaction, which approached significance and suggested a possible trend in the predicted direction for pronoun targets. Since an effect of Distance on reading time had been predicted only for sentences containing pronouns, the data for these target sentences were analysed separately, in a two-way analysis of variance with reader Skill as a between-subjects factor and with Distance as the only within-subjects factor. The ANOVA summary table is given in Appendix 14. In this analysis, however, the main effect of Distance remained nonsignificant, with $F < 1$, while the interaction with Skill was also nonsignificant, with $F(2,44) = 1.68$. It may be noted from the left half of Table 1.3, however, that any possible Distance x Skill interaction clearly does not go in the direction which would be predicted by such arguments as those of Daneman and Carpenter (1980); it was the skilled reader group which showed a trend for reading time to increase with increasing numbers of intervening sentences, while less-skilled readers did not. This would be consistent, however, with the view that the distance effect was caused by confounded topic changes, to which poor readers were less sensitive.

One possible explanation for the failure of the predicted interaction to emerge is that the Nelson-Denny test may have been an unsuccessful measure of reading skill. The absence of a main effect of

Skill on target sentence reading times is one indicator of this; although the two groups were widely separated on reading rate as measured by the Nelson-Denny, this was not reflected in their reading speeds in the main experiment. As a further test of this, reading times for all the sentences in the experimental passages presented to subjects were analysed. Each subject read eighteen passages of six sentences each, giving a total of 108 sentences; when the two skill groups were compared on mean 108-sentence reading time, there was again no difference between them, with $F(1,22)=1.60$. Mean 108-sentence reading time also had nonsignificant product-moment correlations with both Nelson-Denny scores. Finally, it may be noted that the two skill groups also did not differ on comprehension performance in the main experiment; for the comparison between the two groups on accuracy on the eighteen comprehension questions, $F(1,22)=2.18$, which was not significant, while for the comparison on answer latency, $F<1$. It may be noted that overall mean accuracy on the comprehension questions was 16.21, out of a maximum possible of 18.

It thus seems clear that the Nelson-Denny test was a poor predictor of reading performance within the main experiment. It was therefore decided to redivide the subjects into two new skill groups, based solely on their overall sentence reading speed. The twelve subjects with the fastest mean 108-sentence reading times formed the new skilled reader group, while the twelve subjects with the slowest mean 108-sentence reading times formed the new less-skilled reader group. No account was taken of comprehension performance in this division. Although the new division into skill groups still meant that each experimental condition was equally represented in the two groups, it did mean that the groups were no longer balanced in terms of the particular passages read in each condition. Any findings would thus have needed cautious interpretation.

In fact, however, this new rearrangement produced no changes in the pattern of results. The analyses carried out were identical to those described above, except for the new subject division. On the overall analysis, the main effect of Skill was naturally highly significant, with $F(1,22)=14.49$, $p<0.001$, but there were no interactions involving that factor. On the separate analysis of pronoun targets, both the main effect of Distance and the interaction with Skill yielded F values of less than 1. The new skill groups did not differ on comprehension question accuracy, ($F<1$), but did differ on comprehension answer latency ($F(1,22)=16.67$, $p<0.001$). The difference in mean latency was 1565msec; the difference between the two groups on mean target sentence reading time was 1482msec. The difference in answer latency may therefore simply have been due to the time taken to read the question.

The analyses of principal interest thus failed to support the main predictions of the experiment. When the data for pronoun target sentences were analysed separately, reading times were seen to be unaffected by target-antecedent distance; although there was a slight trend in the predicted direction, this was very weak. In addition, the predicted interaction between Distance and Skill on pronoun targets did not appear.

There were, however, two other results of interest, which appeared in a full analysis of all 108 sentence reading times. Firstly, it was found that there was a significant effect on reading times of text serial position. For each subject, total 6-sentence reading times on the texts in sequential positions 1-18 were calculated, and an analysis of variance was carried out on these passage reading times. Different passages in different forms occupied the eighteen sequential positions for different subjects, but the passages were all of similar length, and

this randomisation will have reduced the danger of specific materials effects. It was found that passage reading time was significantly affected by sequential position, with $F(17,374)=4.16$, $p<0.001$. The later passages tended to be read faster than earlier ones; for the first sequential position, mean reading time was 16.507 seconds, while for the last sequential position, mean reading time was 13.719 seconds. This effect was presumably due to practice and familiarity with the task.

The second, and more interesting, finding, was that sentence reading time was affected by serial position within the passages. For each subject, the mean sentence reading times for serial positions 1-6 were calculated over all eighteen passages read, and these figures were analysed; the ANOVA summary table is given in Appendix 14. For the main effect of serial position, $F(5,110)=11.57$, $p<0.001$; sentences late in a passage tended to be read slower than sentences early in a passage. The most interesting aspect of this result, however, was the interaction with reader skill. There was no interaction of serial position with Skill on the Nelson-Denny division, ($F<1$), but on the new skill division as described above this interaction was highly significant, with $F(5,110)=5.79$, $p<0.001$. Table 1.4 presents the mean reading times for serial positions 1-6, for the new skilled and less-skilled reader groups. It can be seen that while the faster readers tended to read through the passages at an approximately steady rate, the slower readers tended to become progressively slower on the later sentences. Slower readers took nearly 1500msec longer to read the sixth sentence than they took to read the first sentence, while for fast readers the difference was only 200msec. It thus appears that discriminating between subjects on the basis of their reading speeds within the experiment did also discriminate between subjects with different reading strategies. The

slower readers were slower at all serial positions, including the first, but the pattern of their reading times was markedly different from that found for the faster readers.

Table 1.4

Experiment 1: Mean Reading Times (msec) for Serial Positions 1-6

	<u>SERIAL POSITION</u>						
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	
<u>NEW</u> <u>SKILLED</u> <u>READERS</u>	1640	1839	1826	1972	2003	1862	1857
<u>NEW LESS</u> <u>SKILLED</u> <u>READERS</u>	2153	2526	2702	3077	3471	3625	2926
	1897	2183	2264	2525	2737	2742	

An interesting methodological problem is presented by the analyses contained in the above two paragraphs. Because these analyses used within-subjects factors with relatively many levels, the degrees of freedom associated with the F ratios were correspondingly inflated; the possible dangers of this are best seen in the analysis of total passage reading times, where a comparatively small F value nonetheless reached a high significance level. A procedure which is often recommended in such cases is the use of conservative degrees of freedom; as described, for example, by Edwards (1972), this procedure involves dividing the full degrees of freedom by a figure which is one less than the number of repeated measures. By this method, the result of the total passage reading time analysis would become $F(1,22)=4.16$, which narrowly fails to achieve significance. In a similar way, the Distance x Reference Type

interaction noted in the analysis of target sentence reading times would not even approach significance if conservative degrees of freedom were used. Clearly, the use of conservative degrees of freedom can greatly alter the obtained significance levels. Since this procedure is a matter of some controversy, and since it is recognised that conservative degrees of freedom are an extremely demanding test of significance, the present research will normally make use of full degrees of freedom; however, in cases where the conservative procedure would lead to a result failing to reach significance, this will continue to be pointed out.

Discussion

Two predictions were made in this experiment. Firstly, based on the theory that pronouns are easier to interpret if their antecedents are active in working memory, it was predicted that pronoun sentence reading times would increase with pronoun-antecedent distance; and secondly, based on the theories reviewed earlier concerning individual differences in reading ability, it was predicted that good and poor readers should be differentially sensitive to this effect. Neither of these predictions was strongly supported by the results. Although pronoun sentence reading times increased by over 300msec from distance 0 to distance 2, this did not prove significant even when the pronoun data were analysed separately; while none of the interactions involving reader skill even approached significance.

The failure to find a significant main effect of distance on pronoun sentence reading times is an interesting one, since this

conflicts with many previous findings such as those of Clark and Sengul (1979). The fact that one improvement in the present experiment should have acted to strengthen and extend the effect found by Clark and Sengul makes the disagreement especially difficult to interpret. In addition to this conflict with previous studies, one other reason for being cautious in concluding that no distance effect exists may be found in the narrowly non-significant interaction of distance with type of reference. This interaction gives some support to the view that anaphoric references and references to new entities are differentially affected by distance; pronouns became harder to interpret, as predicted, while new noun phrases became easier. This latter effect was not predicted, but may also be due to the presence or absence of the antecedent in working memory. It is possible that at short distances the presence of the antecedent in working memory made the sudden introduction of a new actor more difficult, as it involved an abrupt topic change; while at distance 2, with the antecedent no longer in an active state, mention of a new individual may have been less disruptive. A notion related to this will be given fuller discussion in connection with Experiments 8 and 9.

In summary, trends in the data give slight support to the view that pronoun-antecedent distance affects ease of pronoun interpretation, but these trends did not reach full significance. One possible reason for this, and the most obvious difference between this experiment and studies which have found a significant distance effect, lies in the mode of presentation used. It was argued in the introduction that cumulative presentation would allow regressive eye fixations to be made when antecedents were not available in working memory; however, it is also possible that this presentation technique may have encouraged exhaustive regressions even when these were not strictly required. It is

interesting to note that Carpenter and Just (1977a), using a similar presentation technique while recording eye fixations rather than reading times, found that regressions occurred equally often whether or not the pronoun's antecedent was foregrounded. The possibility therefore remains open here that all subjects made full regressive searches in all experimental conditions. This could possibly have swamped any effect of distance on working memory presence. If this is the case, it would suggest that the distance effect, and possibly other types of foregrounding, were of less relevance in normal reading than in the artificial sentence by sentence reading widely used in psychological research.

The lack of any interactions involving reader skill contradicts the findings of Daneman and Carpenter (1980), especially since any trends in the present data in fact showed skilled readers to be more affected by distance than less-skilled readers. However, the absence of skill interactions is made difficult to interpret by the absence of any main effects of reader skill; the Nelson-Denny test did not predict either reading speed or comprehension accuracy within the main experiment, despite that fact that the two skill groups were very widely separated on the screening test. One possible explanation for this is that reading from a computer screen may greatly alter subjects' reading strategies, such that print and computer reading are genuinely uncorrelated. The reading of computer text has become a separate research field in its own right, as the use of computer information systems becomes increasingly widespread (see, for example, papers in Kollers, Wrolstad and Bouma, 1980); one general finding of such research is that presentation mode can have significant effects on reading strategy. This possibility was implicit in the above discussion of the effects of cumulative presentation.

A second explanation for the absence of any main effect of skill is that the Nelson-Denny test may simply be an inexact measure of reading ability. The lack of any correlation between its two performance measures is worrying, while the measurement techniques themselves can be criticised. The speed measure records the number of words read in exactly one minute, and in addition to this small sampling period is dependent on accurate marking by the subjects themselves; while for the comprehension test the questions are in full view next to the material to be read, allowing subjects to adopt a strategy of scanning for the required information. Neither measure is therefore fully convincing, and the failure to predict performance within the main experiment clearly vitiates conclusions regarding skill differences.

The final aspect of the results to be discussed is the effect of sentence serial position. Serial position effects have been frequently reported in the literature, but the nature of the effect varies between studies. Carpenter and Just (1977a) found that later sentences in a passage had longer reading times than earlier ones; this occurred whether subjects were deciding on the consistency of each sentence with what had gone before, or were simply reading normally. Garrod and Sanford (1977) also found a positive effect of serial position; reading times for a movable constant sentence were longer when it appeared in third position in a passage than in second. The present experiment confirmed these findings, with subjects showing significantly longer reading times for later sentences. The explanations offered by Carpenter and Just and by Garrod and Sanford are quite similar, claiming that memory load increases as more material is read.

By contrast with these results, however, there are very many studies that have found a negative effect of serial position. Cirilo

and Foss (1980), for example, found reading times for sentences to be significantly shorter if they occurred in the second half of a story than in the first; this was attributed to the greater amount of new content contained in early sections of a passage, which requires greater processing. Other studies reporting a significant decrease in reading time with serial position include Haberlandt (1980) and Haberlandt and Bingham (1978), who also refer to the additional processing involved in setting up new memory locations early in stories. Graesser, Hoffman and Clark (1980) found a nonsignificant negative relationship between serial position and reading time; however, they used a multiple regression technique to partial out the effect of having more new arguments early in a passage, and agreed with Cirilo and Foss in attributing the serial position effect to the extra processing which these require.

These contradictory results may be reconciled by considering the mode of presentation employed in each case. Carpenter and Just (1977a) employed a cumulative sentence presentation technique very similar to that used here. Garrod and Sanford (1977) are not explicit on this point, but as their apparatus was a teletype rather than a computer screen it appears that in their experiment also the earlier sentences remained visible while later ones were displayed. Studies showing a negative effect of serial position, however, have employed sentence by sentence presentation, with only one display visible at a time. It therefore appears that the nature of the serial position effect is dependent on presentation technique. The positive effect in studies using cumulative presentation may be due to regressive scanning of the earlier material; this further difference between the two presentation modes thus gives some support to the earlier suggestion that exhaustive regressions may have obscured the predicted distance effect. Which, if either, of the two types of serial position effect would be found in

normal reading is unclear.

Aside from again revealing the importance of presentation mode, the serial position effect also reveals a difference in strategy between fast and slow readers. There was no interaction between serial position and reader skill on the Nelson-Denny division, but when subjects were divided into new skill groups on the basis of their mean reading speeds in the experiment itself it was found that only the slower readers were affected by position. The slow readers were at a disadvantage even on the first serial position, but became increasingly slow on later sentences. The fact that the serial position effect is not shown by all subjects makes it even more difficult to interpret; it could be argued that poorer readers are more likely to suffer from a progressive memory overload as they read through a passage, but this explanation of the positive serial position effect seems inadequate given that the effect is reversed under presentation conditions presumably forcing a reliance on memory. The other possibility is that the positive effect, if due to regressive scanning, would be more pronounced among poor readers who make more such regressions. At all events, it is clear that fast and slow readers differ in the pattern of their reading times, as well as in their overall speeds.

In conclusion, this experiment has given marginal support for the original prediction concerning the distance effect on pronoun interpretation, and has also provided evidence that presentation technique can crucially alter reading time results. It has also shown that a standardised reading test may be an unreliable predictor of computer reading, while computer reading speed itself does distinguish different types of reader. These broad conclusions were used in the design of Experiment 2.

Introduction

The primary purpose of Experiment 2 was again to test for the predicted effect of distance on pronoun sentence reading times, and for the interaction of this effect with reader skill. Several significant changes were made from Experiment 1, however, in order to strengthen the likelihood of obtaining these effects, to yield more information about them, and to investigate a number of new questions.

Firstly, the range of pronoun-antecedent distances used was extended up to a gap size of eight intervening sentences. The majority of published studies have used relatively few levels of distance, sometimes having only a near and far case (e.g. Cirilo, 1981); this means that the conclusion by Clark and Sengul (1979), that increasing distance has no further effect beyond a one-sentence gap size, remains untested. It seems unreasonable to assume, even on their proposed discontinuity model, that antecedents not in working memory will be equally easy to retrieve no matter where they appeared in the text. This experiment therefore tested whether further increases in reading time would continue to occur as more intervening sentences were added.

Secondly, this experiment included questions interrogating the referent of the target pronoun, as used by Daneman and Carpenter (1980). In that experiment, the pronoun question technique was found to show progressive increases in pronoun interpretation difficulty as distance was increased up to a gap size of six intervening sentences; this clearly contradicts the suggestion made by Clark and Sengul. This

discrepancy may reflect the fact, noted earlier, that pronoun question accuracy may depend on inferences made at the time of answering, rather than on immediate processing difficulty while reading. This experiment therefore used the pronoun question technique and the reading time technique with the same subjects and on the same passages, in an attempt to discover whether they produce comparable results.

Thirdly, reading ability in this experiment was assessed by means of a simple speed measure incorporated into the main experimental run itself. Reading speed was measured for all subjects, and they were divided into skilled and less-skilled reader groups at the median of these scores. Clearly, this procedure meant that skill groups could not be determined until after the experiment was completed, and thus materials could not be balanced between groups as they were in Experiment 1. Rather, the method used by Daneman and Carpenter (1980) was adopted, with all subjects seeing the same passages in the same experimental conditions. Although this left a danger that specific materials effects would contaminate the results, this could to some extent be countered by having several passages in each condition, and it was felt to be a justifiable risk in order to ensure comparability of the two skill groups.

The fourth major change from Experiment 1 was the inclusion of the reading span test used by Daneman and Carpenter (1980). Although that study showed reading span and a gross measure of reading ability to be significantly correlated, the main results of interest were the interactions of the distance effect with span level itself. Span was included in the present experiment principally to test for a possible interaction with any distance effect shown by the reading time technique.

The final main change in this experiment was the adoption of the more usual sentence by sentence display procedure, rather than the cumulative presentation technique. The possibility remains that the distance effect, and related phenomena, may be an experimental artifact induced by presentation technique; however, this possibility will only be noted here. Investigation of the effects of presentation technique could produce important findings concerning the generality of various psycholinguistic results, but in common with the bulk of the literature on reading this problem will not be considered here. An incidental prediction based on the change of presentation technique was that the serial position effect found in Experiment 1 would be reversed, with subjects becoming faster on later sentences rather than slower.

A number of minor changes from Experiment 1 will be more briefly explained. In that experiment, subjects were found to become significantly faster on later passages; in this experiment, much more practice was given to subjects, in an attempt to lessen the possible obscuring effects of this increase in speed. In addition to practice passages, this experiment also included distractor passages which varied from the constrained structure of the experimental passages. The experimental passages themselves were quite different in character from those used in Experiment 1; subjects in that experiment commented on the very short sentences, with the same number of sentences in each passage. For this experiment, sentences were made longer and more natural, as were the passages as a whole; sentence length and passage length both varied. Distance in this experiment was thus defined solely in terms of the number of intervening sentences between pronoun and antecedent, with no rigid control of word or syllable numbers. A perhaps more significant change was the inclusion of more than one actor in each passage, in order to provide alternative potential pronoun

antecedents. Although it was ensured that the target pronouns were clearly unambiguous, as will be described below, the need to choose between a set of possibilities was expected to make the task more sensitive to the distance effect; inclusion of only one entity in the materials used in Experiment 1 was also intended to avoid the danger of pronoun ambiguity, but may have made the assignment problem too straightforward for a distance effect to appear.

This experiment was thus mainly concerned with issues of pronoun interpretation; however, two subsidiary manipulations were also included, which were concerned more directly with differences between good and poor readers. One important feature of comprehension is that readers quickly forget both the verbatim surface structure of a text and the specific propositions which it expresses, and rely instead on constructive memory representations which may contain more information than the original material. Forgetting of surface structure was demonstrated by Sachs (1967), who presented subjects with short passages and at various points asked them to decide whether test sentences were identical to or different from ones that had actually been presented. Test sentences which were syntactically altered but preserved the meaning of presented sentences were found to be virtually indistinguishable from genuinely old sentences. Forgetting of the specific propositions of a text was demonstrated by Bransford, Barclay and Franks (1972), who found that subjects tended to falsely recognise test sentences which formed a valid inference from the sentences which had actually been presented.

The question of interest here was the possibility that skilled readers might make greater use of such constructive comprehension processes than less-skilled readers. Oakhill (1982) carried out such a

test on seven and eight year old children, and found that while recognition test accuracy was greater for skilled readers if the test sentence did not describe a valid inference from the presented sentences, accuracy was greater for less-skilled readers if the test sentence did describe such an inference. Thus, skilled readers were building elaborated memory representations of the passages, and so were unable to discriminate actually presented sentences from sentences which were merely consistent with the presented sentences. Two tests dealing with this possibility were used here. Firstly, after reading certain passages subjects were presented with test sentences, and were asked to decide whether or not they had actually occurred in the passage. The test sentences were constructed in a similar fashion to those used by Sachs (1967), such that new test sentences either did or did not preserve the meaning of a presented sentence. Secondly, instead of using recognition tests to investigate memory for inferences based on a text, subjects were given questions about passage content which could only be answered correctly if an inference had been drawn. This procedure was similar to one used by Kintsch (1974), who found that questions interrogating an inference were more difficult to answer than those which could be answered on the basis of an explicitly stated fact; this comparison was included in the present experiment, with the possibility of interest being that less-skilled readers would show this difference more strongly than skilled readers.

It should be stressed that these two tests were subsidiary topics in the present experiment, and were intended more as pilot work for possible future studies than as an attempt to give them a rigorous investigation. In particular, it will be seen in the Method section that very few recognition tests were in fact given; this was done in order to minimise any resulting alterations in subjects' reading

strategies which might interfere with the principal effects of interest.

Method

Subjects

24 Plymouth Polytechnic students were recruited through a sign-up notice, and received £2.50 for their participation. There was no pre-test of reading ability; subjects were divided into skilled and less-skilled reader groups only after the experiment was completed. This division will be described below.

Materials

a) Experimental Passages

Twenty passages were written as the main experimental materials. The passages were short stories, varying in length from 12 to 20 sentences, which dealt with topics as diverse as the events at a company board meeting or the experiences of a travelling carnival in turn-of-the-century Europe. Two factors were varied. Each passage included one target sentence, which contained a pronoun that referred to a particular entity introduced earlier in the text. In half the passages the target sentence was the last sentence in the passage, while in the other passages the target was followed by four more sentences. This will be termed the Position factor, and was introduced primarily to conceal the pattern of pronoun occurrences. It was not expected that

Position should have any effect on target sentence reading times. Position might, however, have an effect on pronoun question accuracy, if questions are harder to answer the further back in the passage the relevant information has occurred. The second factor was the distance from the target sentence back to the sentence which had introduced the pronoun's antecedent; there were either 0, 2, 4, 6, or 8 filler sentences intervening between the target sentence and the antecedent sentence. Taken together, the two factors of Position and Distance define ten experimental conditions; two different passages were written for each condition. It should be stressed again that unlike Experiment 1, where all the passages could be varied to fit into each condition, in Experiment 2 each passage was written specifically for one condition, and never varied in form.

In order to ensure comparability across conditions, the target sentences were all of approximately the same length, varying from nine to eleven words. The pronouns used were 'he' (8 times), 'she'(3), 'his'(3), 'her'(3), 'it'(2), and 'him'(1). The position of the pronouns in the target sentences was varied. Antecedents were either proper names or definite noun phrases, which referred to people, animals, or things; the antecedent entities were mentioned only once in each passage, in the antecedent sentence. The sentences preceding the antecedent sentence introduced several other entities of the same type as the antecedent; this was in order to make the choice of pronoun referent more demanding, by providing more than one alternative which was syntactically acceptable on such grounds as gender. The number of such alternatives was varied, while in addition other entities which were not potential antecedents were also included. The pronouns were always unambiguous on semantic or pragmatic grounds, in that the target sentence predicated something of the pronoun which was related to what

had been said about the antecedent in the earlier sentence. The antecedent was always the last-mentioned plausible alternative, and was the only one mentioned in the antecedent sentence. The filler sentences which intervened between the target and antecedent sentences made no mention of any of the specific individuals introduced earlier, but dealt in more general terms with other aspects of the narrative. As far as was possible, the passages were written to seem natural and unforced; to this end, the filler sentences were not rigidly controlled in either length or content.

There were three comprehension questions to each passage. The first was always a Wh-question, asking who or what had performed a certain action or undergone a certain experience; this action or experience was always that which had been ascribed to the pronoun in the target sentence. In order to answer the question, by supplying the name or other identifier of the entity involved, it was thus necessary to have assigned the pronoun correctly; earlier criticism of this technique should, however, be borne in mind. The second question interrogated either a fact explicitly stated in the passage, or an inference derivable from it. These were pragmatic inferences, such as that involved in making the causal connection between a statement that a strong wind was blowing and a later statement that a door flew open. Half the second questions interrogated facts, and half inferences. The third question always interrogated an explicitly stated fact. The location in the passages of the information interrogated by the second and third questions was varied. The questions occasionally required a simple yes or no response, but more usually needed a short phrase or sentence.

Table 2.1

Experiment 2: Example Passage

On the day of the local league snooker final, Frank Worth was taken ill. Harry, Don, Ian, and the other members of his team were frantic. Sam Brayshaw, the league chairman, ruled they could pick a substitute. Only Jeff, a total novice, was available at such short notice. On the face of it, the team's prospects were not very good. The opposing side were very strong, and naturally felt self-confident. Snooker is not an easy game, and it takes time to learn it well. As with all sports, real skill only comes after long practice. The spectators realised the match would now be less entertaining. A few people grumbled as they stood around the playing area. Both teams were nevertheless determined to play a good game. Neither side would admit that the result was a foregone conclusion. He might after all turn out to have a lucky game. The first frame of the match began as the clock struck eight. By the time of the first interval, the scores were level. The conversation around the bar was lively and excited. Everyone was getting caught up in the finely balanced contest.

- 1) Who might turn out to have a lucky game?
- 2) Where did the conversation take place during the interval?
- 3) At what time did the match begin?

An example of the experimental passages is contained in Table 2.1. In this example, the target sentence is fifth last in the passage, the target-antecedent gap is 8 sentences, and the second question interrogates an explicitly stated fact. All twenty experimental passages are given in full in Appendix 2.

b) Distractor Passages and Recognition Tests

In addition to the twenty experimental passages, ten distractor passages were written. These were primarily intended to conceal the patterns in the experimental passages, and were therefore written with a relatively unconstrained structure. They each had three comprehension

questions, none of which interrogated pronouns. The passages are given in full in Appendix 3.

Although primarily intended to provide variation in the structure of the passages read by subjects, the distractors were also used for the sentence recognition tests. This was a subsidiary issue in the experiment, aimed at the possibility that skilled and less-skilled readers might differ in their memory for surface and meaning features of texts, and it was felt desirable to keep this separate from the experimental passages concerned with pronoun interpretation. Test sentences were written for six of the distractor passages. The test sentences were either Olds, which were identical to a sentence in the passage; meaning preserving News, which were different in structure to an actual passage sentence but paraphrased its meaning; or meaning changing News, in which both surface features and semantic content were changed. There were two test sentences of each type, only one test being made on each of the six passages. The New sentences were derived from actual passage sentences principally by changes in word order, but also with some changes in the actual words used. All test sentences were of similar length, ranging from twelve to fifteen words; the sentences are given in full in Appendix 3.

Fourteen practice passages were also written. These were similar to the distractors in being of a relatively free structure, but in some cases the first of the three questions did interrogate a pronoun. Two of these practice passages were followed by recognition tests, both being meaning preserving News. The practice passages and the two test sentences are given in Appendix 3.

c) Sequence Construction

Two random sequences were constructed from the 44 passages. First, half the practice passages were randomly assigned to one sequence, and half to the other, with the constraint that there should be one practice recognition test in each sequence. The seven practice passages were arranged in random order at the start of each sequence. Second, one passage from each experimental condition was randomly assigned to each sequence. Similarly, half the distractor passages were randomly assigned to each sequence, with the constraint that there should be one of each type of recognition test in each sequence. The experimental and distractor passages assigned to each sequence were then randomly mixed together. Each sequence thus consisted of seven practice passages followed by a random arrangement of ten experimental and five distractor passages, making 22 passages in all. Both sequences were presented to every subject.

d) Reading Speed Passages

Subjects were divided into skill groups solely on the basis of their reading speeds, and two additional passages were written especially for this purpose. These passages were completely free of experiment-imposed constraints, and were written to be as natural as possible. One of the passages described three friends on a walking trip, while the other was set in a snowbound school in winter. The first passage was 321 words long, and the second was 315 words long. Each passage was followed by three questions, all of which interrogated explicitly stated facts. One of these passages was randomly assigned to each sequence. The reading speed passages are given in Appendix 4.

e) Span Test Sentences

The materials for the reading span test were constructed in exactly the manner described by Daneman and Carpenter (1980). Sixty sentences, occasionally requiring slight modification, were taken from novels and short stories. The sentences varied in length from thirteen to sixteen words, and each one ended in a different word. The final words were usually singular common nouns, but there were also some plurals and adjectives. Each sentence was typed on a single line across the centre of an 8x5in white index card. The sentences were randomly grouped into sets, which contained either 2, 3, 4, 5, or 6 sentences; there were three sets of each length. The cards were arranged in a pack, beginning with the three sets of length 2, followed by those of length 3, and so on. Blank cards were inserted between the sets. There were also ten practice sentences, randomly arranged into five sets of two sentences each. The sentences used are contained in Appendix 5.

Procedure

Because of the large number of passages that were to be read, it was necessary to conduct the experiment over two separate one-hour sessions, with one of the 22-text sequences being presented in each. Each subject saw the two sequences in the same order. As with the fact that all subjects saw the same passages in the same conditions, this was done to allow a post-experiment skill division into equivalent groups. Some subjects went through the two sessions with only a five-minute break in between, while for other subjects, who were unable to attend

for two consecutive hours, there was an interval of up to a week. Testing arrangements were identical to those described for Experiment 1, using the same computer system and the same general instructions.

As before, subjects began each passage by pressing a red button when the warning message was displayed. The passages were then presented one sentence at a time, with each sentence appearing centred on the middle line of the screen. After each button press, unlike Experiment 1, the current sentence was cleared from the screen before the next was displayed. Reading times for each sentence were again recorded automatically. When the button was pressed after reading the last sentence of each passage, it was replaced by the three questions, which were displayed on the screen together. Subjects wrote their answers in a booklet. They were told to decide for themselves what constituted a complete and appropriate answer in each particular case, and to give whatever information they felt was relevant and necessary. If they were unsure on any question, they were encouraged to make an attempt. The questions remained on the screen until subjects had written their answers and pressed the red button to proceed.

Pressing the red button after writing the question answers normally caused the warning message for the next passage to be displayed. However, on those occasions when there was to be a recognition test, it led to the questions being replaced by the test sentence. This appeared across the middle of the screen exactly as the passage sentences had done; simultaneously, a prompt message saying 'Did this sentence appear in the previous text?' was displayed at the top of the screen. This ensured that subjects did not confuse test sentences with the start of a new passage. Subjects were warned in the initial instructions that recognition tests would occur very occasionally, and the format of the

tests was described to them. They were told to decide whether or not the exact sentence had appeared in the previous passage. To give their decision they pressed one of two buttons on the right of the response-key board, labelled respectively 'YES' and 'NO'. Subjects were told to guess if they were unsure. When the response had been given, the test sentence and prompt disappeared from the screen, and in two seconds were replaced by the warning message for the next passage. Recognition test answers and response latencies were recorded automatically.

Subjects read through all 22 passages in each sequence in this way, and were not told that the first seven passages were in fact only for practice. Instructions were given orally at the beginning of the first session, when subjects also read through two additional short passages, in the presence of the experimenter, as explicit practice. The second of these practice passages had a recognition test, and both passages are given in Appendix 3. At the beginning of the second session, subjects were simply told that it was identical to the first, though with different passages. Subjects were left alone to read through the two sequences.

In each session, the reading speed passage that had been assigned to that sequence was interpolated after the fourth concealed practice passage. Subjects were warned that one passage, early in the sequence, would be displayed in its entirety, filling the whole screen, and that in that case they should not press the red button until they had read the complete text. Because of machine constraints the full passages could not be displayed instantaneously, but instead were displayed line by line, in rapid succession, from the top of the screen to the bottom. Reading time was measured from the moment the text began to be

displayed, as it would be possible for subjects to read the first few words while the rest of the passage was being printed on the screen. Both passages took up 23 lines of the 24-line screen, and were arranged as single paragraphs. When the button was pressed after completing the passage, the questions were displayed in the usual way. The reading speed passages came fifth in the sequence so that subjects would read through three more sentence by sentence practice passages before the experimental passages began. Any disruption caused by the transition to and from full-screen reading was therefore minimised.

The reading span test was given at the end of the second session. The cards were placed one by one on a desk in front of the subject, who read the sentences out loud. As soon as subjects had finished reading one sentence, the next card was placed on top of it. When the next card was blank, indicating the end of a sentence set, subjects attempted to repeat back the last word from each of the sentences in that set, in the order in which they had been presented. Subjects were told to read the sentences aloud at their own pace. They were also told that the sentences were completely unrelated to one another. After the five two-sentence practice sets had been attempted, a tape recorder was switched on before commencing the experimental sets. Subjects were warned that the number of sentences in each set would increase through the experimental pack, but were not told that there were three sets of each length. Testing continued until the subject failed all three sets at any one length, and was then terminated.

Results

The analyses to be described first are those conducted with subjects divided into skill groups on the basis of their reading speeds; the outcome of the reading span test will be described towards the end of this section.

Main Analyses

For the reading speed division, subjects were discriminated solely on their reading times for the two full-screen passages. Reading speeds on these two passages had a product-moment correlation of +0.911 ($df=22$, $p<0.005$), and the skill division was therefore made on the mean of these two measures. The twelve subjects with the fastest mean full-screen reading speeds formed the skilled reader group, while the twelve subjects with the slowest mean reading speeds formed the less-skilled reader group. As each subject had read exactly the same materials, any subject could be assigned to either group without affecting their comparability.

Table 2.2 presents the mean target sentence reading times in milliseconds for the ten experimental conditions. The data were analysed in a $2 \times 2 \times 5$ analysis of variance, with Skill as a between-subjects factor and Position and Distance as within-subjects factors. As each subject read two passages in each condition, the analysis was conducted on the mean of these two values. The ANOVA summary table for this analysis is given in Appendix 15.

Table 2.2

Experiment 2: Mean Target Sentence Reading Times (msec)

<u>POSITION</u>	<u>DISTANCE</u>					
	<u>0</u>	<u>2</u>	<u>4</u>	<u>6</u>	<u>8</u>	
<u>LAST</u>	2067	2557	1942	3071	2133	2354
<u>FIFTH LAST</u>	2020	2569	3269	3228	2722	2761
	2044	2563	2605	3150	2427	

It was found that there was a significant main effect of Skill, with $F(1,22)=12.80$, $p<0.005$; mean reading time for skilled readers was 1997msec, while mean reading time for less-skilled readers was 3118msec. This difference of over 1100msec shows that full-screen reading speed successfully predicted sentence by sentence reading speed. There were, however, no interesting interactions involving the Skill factor, and means within the two groups are therefore not presented. For the Distance x Skill interaction, $F(4,88)=1.09$, while for the Position x Skill interaction, $F<1$. There was a significant three-way interaction of Distance x Position x Skill, with $F(4,88)=3.22$, $p<0.025$, but this would not be significant with conservative degrees of freedom and seemed to be due primarily to some very long reading times in the less-skilled group at distance 4 and in fifth last position. As a check on this, the analysis was re-run with all extreme reading times omitted. For each subject, the mean and standard deviation of their target sentence reading times were calculated, and any observation more than two standard deviations from the individual's mean was deleted. When the analysis was re-run using the GENSTAT statistical package which estimated values for the missing observations, the F ratio for the

three-way interaction was reduced to less than 1; none of the other results to be described were greatly altered. Skill therefore does not appear to have affected sensitivity to the other experimental effects.

The results concerning the other factors in the experiment will now be described. It was found that there was a significant main effect of Distance, with $F(4,88)=7.51$, $p<0.001$. From Table 2.2 it can be seen that reading times increased from distance 0 to distance 2, and from 4 to 6, but that they then fell from 6 to 8. These differences were tested using the Newman-Keuls procedure, and all three comparisons were found to be significant. For the 0-2 comparison, $Q(3)=3.57$, $p<0.05$; for the 4-6 comparison, $Q(2)=3.75$, $p<0.01$; and for the 6-8 comparison, $Q(4)=4.98$, $p<0.01$ (all $df=88$). The comparison between 0 and 8 was not significant. It therefore seems that although the predicted distance effect was present, it had a quite irregular form. The increase in reading times from 0 to 2 constitutes the basic predicted effect, but the second increase from 4 to 6 was surprising given that 2 and 4 were almost equal. A quite unexpected result was the fall in reading times from 6 to 8, with no significant difference between 0 and 8.

There was also a significant main effect of Position, with $F(1,22)=11.46$, $p<0.005$. From Table 2.2 it can be seen that target sentences appearing last in a passage were read faster than target sentences which were followed by four more sentences. This was an unexpected result, as the Position factor had been included primarily to conceal the pattern of pronoun usage in the experimental materials, and it had not been predicted that it should affect reading times. Subjects had no warning that the end of a passage was approaching. The effect is made even more difficult to interpret by the fact that it appeared only

at distances 4 and 8. When the simple main effects of Position were tested separately at the five levels of Distance, F values of less than 1 were obtained at 0, 2, and 6; while at 4, $F(5,110)=31.01$, $p<0.001$, and at 8, $F(5,110)=6.24$, $p<0.001$. These variations in the effect of Position were reflected in the significant Distance x Position interaction, with $F(4,88)=6.12$, $p<0.001$.

The reading time results were closely paralleled by the results of the analysis of pronoun question accuracy. Table 2.3 presents the mean number of correct answers to the questions which interrogated the referent of the pronoun in the target sentence; as each subject read two passages in each condition, the maximum possible in each cell is 2. These data were analysed as a 2 x 2 x 5 design, exactly as described for target sentence reading times; the ANOVA summary table is given in Appendix 15. There was no main effect of Skill ($F<1$), and no interactions involving that factor (for Distance x Skill and Distance x Position x Skill, $F<1$; and for Position x Skill, $F(1,22)=2.55$). There was, however, a significant main effect of Distance, with $F(4,88)=9.13$, $p<0.001$. From Table 2.3 it can be seen that accuracy fell from 0 to 2, and again from 4 to 6, but that it increased from 6 to 8. Newman-Keuls tests showed all these comparisons to be significant. For the 0-2 comparison, $Q(4)=4.66$, $p<0.01$; for the 4-6 comparison, $Q(3)=4.66$, $p<0.01$; and for the 6-8 comparison, $Q(4)=4.66$, $p<0.01$. The difference between 0 and 8 was also significant, with $Q(2)=3.84$, $p<0.01$. (All $df=88$).

Table 2.3

Experiment 2: Mean Pronoun Question Accuracy (max=2)

<u>POSITION</u>	<u>DISTANCE</u>					
	<u>0</u>	<u>2</u>	<u>4</u>	<u>6</u>	<u>8</u>	
<u>LAST</u>	2.000	1.708	1.625	0.958	1.250	1.508
<u>FIFTH LAST</u>	1.500	1.083	1.292	1.250	1.667	1.358
	1.750	1.396	1.458	1.104	1.458	

The effect of Distance on pronoun question accuracy was thus an almost perfect reflection of the effect on target sentence reading times; increases in reading times were matched by decreases in accuracy, while a fall in reading times from 6 to 8 was matched by an increase in accuracy. The congruence between these two sets of results was predicted on the assumption that both would be affected by the degree of difficulty of pronoun interpretation. Although the 0-8 comparison was not significant for reading times, the trend was clearly in the same direction as the effect on accuracy. By comparing Table 2.2 and Table 2.3 it may be seen that the effect of Position was also apparently present on both measures; reading times were longer when the target sentence was followed by four more sentences, and accuracy on pronoun questions was correspondingly lower. In the analysis of variance, the main effect of Position on pronoun question accuracy in fact marginally failed to reach significance, with $F(1,22)=4.22$, $p<0.1$; however, the correspondence between the reading time and accuracy results was nonetheless a close one.

There was also a significant Distance x Position interaction on pronoun question accuracy, with $F(4,88)=7.98$, $p<0.001$. In this case, however, the results on accuracy were not in complete agreement with those on reading times. Comparison of Table 2.2 and Table 2.3 shows that the effect of Position was not the same at each distance for the two measures. The effect of Position was only significant at distances 4 and 8 on reading times, while for accuracy it can be seen that the effect was in fact reversed at distance 8, and was strongest at distance 2. Given the unexpectedness of even the main effects of Position, the Distance x Position interactions are essentially uninterpretable.

Subsidiary Analyses

The above analyses were the ones of principal interest to the experiment, dealing with the effects of target-antecedent distance on the ease of pronoun interpretation. The results of the other features of the experiment will now be reported, somewhat more briefly.

Accuracy scores on the second question were analysed as a 2 x 2 design, with Skill as a between-subjects factor and Question Type (fact or inference) as a within-subjects factor. The ANOVA summary table for this analysis is given in Appendix 15. Questions interrogating explicitly stated facts were found to be answered more accurately than questions requiring inferences; for the former, mean accuracy (out of a possible 10) was 8.96, while for the latter mean accuracy was 8.08, with $F(1,22)=7.84$, $p<0.025$. The main effect of Skill and the interaction were not significant (both $F<1$). Accuracy on the third question, which always interrogated an explicitly stated fact, was analysed as a 2 x 2 x 5 design, with Skill, Position and Distance as factors; it was not

expected that experimental condition should have any effect on accuracy, and it was mainly of interest to discover whether the two skill groups differed in comprehension performance as well as in speed of reading. As it turned out, the only significant effect in this analysis was the main effect of Distance, with $F(4,88)=2.87$, $p<0.05$; it was noticeable that while accuracy was high at distance 0, it was relatively lower at all the other distances. This effect can presumably be attributed to specific materials differences, and would in any case be non-significant with conservative degrees of freedom. The main effect of Skill was not significant, with $F(1,22)=1.28$; mean accuracy overall was 1.76, out of a maximum possible of 2.

Recognition test latencies were analysed as a 2 x 3 design, with Skill as a between-subjects factor, and Test Type (Olds, meaning preserving News, and meaning changing News) as a within-subjects factor. The ANOVA summary table is contained in Appendix 15. There was a main effect of Skill, with $F(1,22)=5.64$, $p<0.05$; mean latency for skilled readers was 5749msec, while mean latency for less-skilled readers was 8028msec. The main effect of Test Type and the interaction were not significant (both $F<1$). Recognition test accuracy scores were analysed in a similar way, and the summary table for this analysis is also given in Appendix 15; the main effect of Skill and the interaction were not significant (for the main effect, $F<1$, and for the interaction, $F(2,44)=1.48$). There was, however, a significant main effect of Test Type, with $F(2,44)=25.87$, $p<0.001$. Mean accuracy (maximum=2) was 1.5 for Olds, 0.46 for meaning-preserving News, and 1.33 for meaning-changing News. Newman-Keuls tests showed that accuracy scores for both Olds and meaning-changing News were significantly greater than for meaning-preserving News, but that they did not differ from each other. For the comparison of meaning-preserving News with Olds,

$Q(3)=13.36$, $p<0.01$, and for the comparison with meaning-changing News, $Q(2)=11.22$, $p<0.01$ (both $df=44$).

The above analyses were conducted with subjects divided into skill groups on the basis of their reading speeds, and the lack of interactions with this factor indicates that the two groups were not differentially sensitive to any of the reported effects. The results of the reading span test, and of the analyses with subjects divided into groups on the basis of their span level, will now be reported.

Each subject's span level was defined as the longest set length on which they scored correctly on at least two out of the three sets. If a subject scored correctly on two out of three at one level, and then on only one out of three at the next level, the span was defined as the former length. In this way, thirteen subjects were found to have spans of 2 sentences, eight were found to have spans of 3 sentences, and only three had spans of 4 sentences. No subject scored correctly on any set at length 5; performance thus seems to have been at a lower level than that reported by Daneman and Carpenter (1980). Span level had a product-moment correlation of $+0.317$ with mean full-screen reading speed, which was not significant. For the purposes of the analysis, subjects were divided into two span groups, composed respectively of those subjects with span level 2, and those with spans of either 3 or 4; it was considered that there were too few subjects with spans of 4 sentences to allow them to be treated as a separate group.

All the analyses described above were re-run with subjects divided into the two span level groups; the results obtained will be briefly summarised. The only change in the pattern of results was that the two groups did not differ on either target sentence reading time or

recognition test answer latency, and nor was there a significant Distance x Position x Span interaction on the analysis of target sentence reading times; all other effects remained as before, with slightly different F values but with unchanged significance levels. There were no significant interactions involving the span level factor on any of the analyses.

The final analysis to be reported concerns the sentence serial position effect. In Experiment 1 it was found that sentences later in a passage were read slower than sentences early in a passage, but that this effect was shown principally by the slower readers. It was of interest to determine whether a similar pattern occurred in the present experiment, given that presentation technique had been changed. To this end, six passages were selected for analysis, consisting of one concealed practice and two distractor passages from each sequence. Experimental passages were not included in the analysis because of their relatively constrained structure. The passages chosen for analysis were selected because of their length, the shortest having 17 sentences and the longest 19; these were the longest practice and distractor passages available. Only the first seventeen sentences of each passage were included in the analysis, in order to ensure use of an equal number of sentences at each serial position. Obviously, the length of the sentences at each position varied, but this variation was a presumably random source of error, given that there were six sentences at each position.

The data were analysed as a 2 x 17 design, with subjects divided into skill groups on the basis of their reading speeds; Skill was thus between-subjects, and sentence serial position was within-subjects. There was a significant main effect of Skill, with $F(1,22)=9.19$,

$p < 0.025$; skilled readers had a mean sentence reading time almost a second faster than that for less-skilled readers, thus confirming the result found on the separate analysis of target sentence reading times. There was also a main effect of sentence serial position, with $F(16,352) = 25.97$, $p < 0.001$. The most striking feature of this effect was the extremely long reading time for the first sentence of the passages. Mean reading time for the first serial position was 5832msec, while for the second sentence it was only 3461msec. Reading times then fell irregularly to the seventh serial position, which had a mean reading time of 2532msec. From the seventh sentence onward, reading times fluctuated between 2500msec and 3000msec, with no further evidence of a sustained fall or rise. In brief, subjects began the passages slowly, and then increased their reading speed until after half a dozen sentences times became stabilised. Although there was a statistically significant Skill \times Serial Position interaction with $F(16,352) = 2.49$, $p < 0.005$, it would clearly be non-significant with conservative degrees of freedom and in any case had no obvious pattern, seeming to be due to occasional discrepancies between the two groups in the direction of change between adjacent positions. The overall trend was similar for both groups. The effect of serial position found in Experiment 1 was therefore reversed in the present experiment, while the interaction of serial position with reader skill was also not repeated.

Discussion

The results clearly supported the prediction that pronoun interpretation would be relatively easy if pronoun and antecedent were adjacent, but relatively difficult if additional sentences intervened.

Reading times were fastest at distance 0, and increased significantly at distance 2. Furthermore, there was a second increase in reading times from distance 4 to 6, suggesting that increasing the gap size leads to still more processing difficulty; this contradicts the theory of Clark and Sengul (1979). The reading time results were closely paralleled by the pronoun question results, indicating that these techniques may reflect similar processes. However, the irregular, stepped nature of the distance effect, and in particular the fall in reading times and increase in accuracy at distance 8, suggest some complexities in the processes underlying these results; the interpretation of the distance effect will therefore be deferred to a separate section at the end of this discussion.

The second main prediction of this experiment, that the distance effect would differ between poor readers and good readers, was not supported; skill and distance showed no interaction, on either the reading time results or the pronoun question results. Unlike Experiment 1, this failure cannot be attributed to an ineffective skill division; full-screen reading proved a highly reliable measure of reading speed, and the main effect of reader skill on target sentence reading times was also highly significant. In addition, the two skill groups did not differ in comprehension accuracy on the non-pronoun questions, confirming that the faster readers were being more efficient overall. This experiment therefore suggests that good and poor readers do not differ in sensitivity to the distance effect.

There are clear contradictions between this experiment and the findings of Daneman and Carpenter (1980), especially with regard to the effects of span level. While in their experiment span level and a gross measure of reading ability were found to be correlated, in the present

experiment span level and reading speed did not correlate. More importantly, reading span did not interact with the distance effect, on either question accuracy or reading times; there was no evidence that high span readers were less sensitive to the distance effect than low span readers. There is no obvious explanation for this failure to replicate, other than the fact that the span levels obtained here were rather lower than those obtained by Daneman and Carpenter; in their first experiment, a total of six subjects obtained span levels of 4 and 5, while here only three subjects reached level 4, and none reached level 5. Since in the Daneman and Carpenter experiment it was the level 5 subjects who were the only ones to be unaffected by distance, the absence of a span by distance interaction in the present experiment may be due to the absence of any high span subjects in the sample. It should be noted that Baddeley (1981) claims to have replicated certain aspects of the Daneman and Carpenter results, but gives no details.

To some extent, the validity of the above results, especially of the distance effect, is placed in some doubt by the effect of target sentence position. This factor was included in the experiment mainly in order to conceal the patterns of pronoun occurrences; although it was predicted that pronoun questions would be answered more accurately if the relevant information immediately preceded the question, rather than occurring further back in the passage, no effect on reading times was expected. In fact, the reading time and accuracy results reflected each other with regard to the effect of position just as they did with regard to distance; accuracy was greater with target sentences in last position, and reading times for such targets were correspondingly faster. Although distance and position also interacted on both measures, however, the two interactions took different forms. Since the main effect of target position on reading times, and both interactions,

were quite unexpected, they suggest that other influences were at work than those of direct experimental interest. The most obvious possibility is that this pattern of results is materials specific, and arose because experimental passages were not balanced across conditions. The danger of such specific effects was acknowledged in the introduction, and must be borne in mind when considering other aspects of the present results.

Having sounded this warning concerning the interpretation of the main results of interest, the more subsidiary aspects of the experiment will now be discussed. The findings regarding sentence recognition memory, that old sentences and meaning-changing new sentences are more accurately identified than meaning-preserving new sentences, are in agreement with previous research such as that by Sachs (1967). The lack of any effect on recognition latencies suggests that the inaccuracy concerning meaning-preserving new sentences was not accompanied by greater difficulty in making the decision; false recognitions to these sentences were made just as rapidly as correct recognitions of old sentences, giving further evidence of subjects' inability to distinguish between them. However, the main question of interest here was whether these effects would interact with reader skill; no evidence for this was found. This indicates that skilled and less-skilled readers show equal forgetting of the surface form of the presented sentences, and retain memory only for the semantic content expressed.

The results of the analysis of fact and inference question accuracy are also in agreement with previous findings. Kintsch (1974) found inference questions to be more difficult to answer immediately after reading a passage, and a similar result was found here. More important for present purposes was the fact that skilled and less-skilled readers

showed no difference in sensitivity to the fact-inference manipulation. This to some extent conflicts with the results of Oakhill (1982), but the many differences between that experiment and the present one, such as in the age of subjects and the type of test used, can presumably account for this disagreement. At all events, the conclusion based on both memory tests used here is that skilled and less-skilled young adult readers do not differ in their relative use of constructive memory representations. It should also be pointed out that neither test showed any main effect of reader skill; skilled readers did not have generally superior memory compared to less-skilled readers.

The remaining result to be discussed concerns the effect of sentence serial position. As predicted, the serial position effect found in Experiment 1 was reversed, with subjects becoming significantly faster on later sentences. This agrees with the usual effect found in studies employing sentence by sentence display, and suggests that the positive serial position effect found in Experiment 1 was caused by the cumulative presentation technique. The pattern of the serial position effect found here seems a sensible one. The first sentence position had an extremely long reading time, which presumably reflects the importance of that sentence in establishing the topic and context of the passage; Kieras (1981) also found an exceptionally long first sentence reading time, when his subjects were making explicit topic choices for each passage. The gradual decline in reading times, to an eventually stable level, may be explained either by the reduced amount of new information contained in later sentences, or by the gradual establishment of an integrated passage representation to which later sentences can be more easily connected. There is no way to separate possible alternative explanations here, and this seems generally to be the case in the literature. The responsiveness of the serial position effect to

presentation technique suggests it may be a topic worthy of investigation in its own right. Finally, it should be pointed out that the serial position effect cannot be used to explain the effect of target sentence position on reading times. Targets which were the last sentence of a passage did tend to have later serial positions than targets occurring further back in a passage, with mean serial positions of 14.4 and 12.5 respectively; however, this difference is very small, and by that time the serial position effect had in any case already levelled off.

The Interpretation of the Distance Effect

The reasons for predicting an effect of distance on pronoun sentence reading times, both in the present set of experiments and in earlier published studies, are based on the theory which links pronoun interpretation to the use of working memory in reading. To briefly review the model which has been proposed in the literature, it is suggested that pronoun interpretation involves an attempt to make a direct match to working memory as the first processing stage; subsequent attempts to retrieve or infer an antecedent will occur only if this initial stage fails. Absence of an appropriate antecedent in working memory thus leads to added processing, which is predicted to cause slower reading times; while increasing pronoun-antecedent distance is said to make it less likely that the correct antecedent will in fact be directly available. There are two ways in which this distance effect could occur. On one view, items may be lost from a limited capacity working memory through general processes of decay and displacement; as more text is read and processed, earlier items disappear from working memory with increasing probability. The second

view argues that items are cleared from working memory by strategy-based control processes which actively select some entities and discard others; Clark and Sengul (1979) suggested that entities are cleared from working memory in an all-or-none fashion at sentence or clause boundaries. This view therefore sees working memory loss as the result of text-specific features of processing, rather than as the simple consequence of general processing limitations.

These two theories of the distance effect make conflicting predictions. The first would predict that difficulty of pronoun interpretation should increase linearly with distance, though possibly reaching an eventual asymptote; while the second predicts a single abrupt increase in processing difficulty, at the point of working memory clearance. Neither theory, however, can readily account for the results found in Experiment 2. The distance effect found here was neither linear, nor characterised by a single discontinuity; rather, reading times were related to distance in a step-like fashion, with two marked but separate increases in reading time. These increases were then followed by a large fall in reading time at the longest distance used. It is difficult to account for this pattern on either of the two mechanisms proposed.

One problem which both these accounts overlook is that antecedent retrieval from long-term memory may itself show gradations of difficulty related to distance. Even if the linear effect predicted by the first mechanism had been found to obtain, this would have been inconclusive; rather than being due to increasing probability of absence from working memory, it could be explained by linearly increasing difficulty of long-term retrieval. In a similar fashion, no single reading time discontinuity need exist even if the mechanism favouring active working

memory clearance were the correct account; further increases could occur due to difficulty of retrieval. In general, the processes of antecedent retrieval, and the influences upon them, are a neglected area of research. However, it can be seen that the inclusion of varying retrieval processes within the model of pronoun assignment would make the first section of the Experiment 2 distance effect equally compatible with either account; but even this addition could hardly be made to explain the fall in reading times at distance 8.

Rather than pursue issues of long-term retrieval at this point, an alternative explanation for the irregular distance effect will be considered. The discussion so far has assumed that distance is the sole influence on working memory presence; however, as the earlier review demonstrated, many other influences on ease of pronoun interpretation also exist. The general model under consideration here argues that items may be present in working memory not only because they have recently been mentioned, but also because of their thematic importance; and that clearance from working memory may occur in response to topical discontinuities cued by signals in the text. The previous two experiments, in line with the common research strategy, attempted to study the effects of one suggested factor in isolation from the rest; but the possibility remains that some of the other factors may also have been at work, leading to variations in the foregrounded status of the crucial antecedents, independently of the distance manipulation. A similar danger was acknowledged by Daneman and Carpenter (1980), who say they attempted to make their crucial antecedents approximately equal in thematic importance as this could affect the length of time they remained foregrounded; no test of the successfulness of this attempt was reported. Since the distance effect found here was much more irregular than that found by Daneman and Carpenter, and was

correspondingly more difficult to explain on the basis of existing theory, it was decided to make some attempt to assess the degree of importance or relevance of each crucial antecedent at the point immediately preceding the pronoun sentence. The question at issue was whether variations in antecedent importance could account for the obtained pattern of reading times.

Table 2.4

Experiment 2: Foregrounding Rating Instructions

You are to judge the status of the referent of the critical pronoun, at the end of the sentence immediately preceding the pronoun sentence - i.e. before the pronoun sentence is read.

At that point, consider how important the referent is to the current text, and how closely related it is to the actions or events that are being described. Consider whether the referent is an entity which might naturally have been talked about again at that point, or whether its explicit reintroduction to the text seems unexpected. Consider whether it is an entity which might have continued to be thought about in connection with the current material, or whether it is quite unconnected.

Considering all this, form a single judgment of the salience of the referent at that point in the text. Express your judgment as a number from 1 to 7, with 1 representing extremely high salience, and 7 representing extremely low salience.

In order to assess the status of the crucial antecedents, three sets of salience ratings were obtained for each experimental passage. These ratings were given independently by the experimenter and his two thesis supervisors; clearly, all those involved were aware of the nature of the experimental passages, and of the hypothesis under investigation. Each passage was given a single rating between 1 and 7, according to how salient the antecedent was judged to be immediately before the pronoun sentence was read; a rating of 1 indicated the

highest degree of foregrounding, while a rating of 7 indicated the highest degree of backgrounding. Thus, if variations in antecedent importance were linked to pronoun sentence reading times, the two measures would be predicted to show a positive correlation. To encourage uniformity in the manner of arriving at the salience ratings, a set of instructions was prepared; these are given in full in Table 2.4. Although it can be seen that full understanding of these instructions depended on previous knowledge of the materials being assessed, the second paragraph was made as explicit as possible in its definition of the property of the antecedents which was to be rated.

Table 2.5

Experiment 2: Mean Foregrounding Ratings

<u>POSITION</u>	<u>DISTANCE</u>					
	<u>0</u>	<u>2</u>	<u>4</u>	<u>6</u>	<u>8</u>	
<u>LAST</u>	1.34	2.84	3.83	6.33	4.17	3.69
<u>FIFTH LAST</u>	1.50	3.67	3.83	5.34	3.50	3.57
	1.42	3.25	3.83	5.83	3.83	

The mean ratings for the ten experimental conditions are given in Table 2.5. It was found that mean salience ratings and mean pronoun sentence reading times for the twenty experimental passages were significantly correlated, with a product-moment coefficient of +0.49 (df=18, $p<0.025$); longer reading times were therefore associated with the more backgrounded antecedents. Probably the most striking aspect of the results was the close correspondence between the pattern of salience ratings and the pattern of the distance effect. Ratings were lowest at

distance 0, indicating that antecedents were considered very salient immediately after being first mentioned, but increased substantially at distance 2. There was then little change between distances 2 and 4, but a second large increase at distance 6; while ratings fell back again at distance 8, to a level equal to that at distance 4. These changes exactly reflect the pattern of reading times. It should be pointed out, however, that there was no such close correspondence with regard to the effect of position; the difference between these two mean ratings was very slight, and was in the wrong direction to account for the faster reading times for targets coming last in a passage.

It therefore appears that the irregular pattern of reading times, with regard to the distance effect at least, may be due to variations in the foregrounded status of the antecedents. Although the procedure described above clearly does not constitute a rigorous test of this hypothesis, the danger it reveals is a plausible explanation and deserves further consideration; even though the ratings in themselves do not reveal the source of the foregrounding in any particular passage, the possibility that other factors in addition to distance were at work is an important one.

One of the questions raised by these results concerns the relationship between pronoun-antecedent distance and the antecedent's current topical importance, or foregrounding. The pattern of the ratings suggests that distance and foregrounding may be completely uncorrelated, since even at the longest distance it was possible for antecedents to receive high foregrounding values. This leaves it unclear whether distance itself has any effect on an item's presence in working memory. Previous theorising has tended to assume that distance was simply one more factor which contributed to an item's foregrounded

status; the possibility being raised here is that distance and foregrounding may be independent factors, and that distance alone does not influence working memory presence. This possibility is left open by the ratings results, since the Experiment 2 effects could be explained by the independent variations in foregrounding alone. It is, however, also possible that distance does have some residual effect on ease of pronoun interpretation, but that this was obscured by the other factors. In general, the problem of obscuring variations in foregrounding prevents any conclusions being drawn about the independent effect of distance itself.

The second question left open by these results concerns the effect of reader skill. Since less-skilled readers are claimed to have smaller effective working memory capacity in reading, they should clearly be more sensitive to any distance effect based on gradual loss from working memory due to capacity limitations; however, as noted earlier the evidence that poor readers are less able to identify discourse topics suggests that they should be less sensitive to foregrounding effects based on changes of topical status. The absence of significant skill interactions in Experiments 1 and 2 may be partly due to the counteracting operation of both these factors.

In summary, this review of the distance effect has shown that it cannot unequivocally be attributed to the influence of distance itself. This was demonstrated in the present experiment by obtaining explicit foregrounding ratings for each passage, but the danger applies equally to the results of Daneman and Carpenter (1980). However, the problem does not exist solely in experiments in which different passages were used in the different distance conditions; as noted earlier, it seems likely that distance and topical discontinuity are inherently

confounded, such that even in experiments with balanced materials sets the existence of a pure distance effect cannot be guaranteed. The principal question left open by this experiment is therefore whether distance itself makes any contribution to the difficulty of pronoun interpretation, over and above the effect due to working memory selection based on textual cues; this question was addressed by Experiment 3.

Introduction

The discussion of Experiment 2 argued that there was no conclusive evidence to show that simple textual distance has an effect on ease of pronoun interpretation, and that obtained effects may have been due to variations in logically independent though normally confounded features of topical structure. While such confounded topic changes were expected to produce a discontinuous distance effect rather than a linear one, the fact that topicality varied in unintended ways between the Experiment 2 passages meant that the relationship between distance and foregrounding was obscured. The present experiment therefore used passages in which the two factors of distance and antecedent foregrounding by topical continuity were varied orthogonally; subjects read materials at the different levels of distance with the crucial antecedents either foregrounded or backgrounded. The distance manipulation was straightforward, consisting as before of different numbers of sentences intervening between pronoun and antecedent; the foregrounding manipulation, however, requires slightly more comment.

As noted in the literature review, many suggestions have been made concerning possible influences on foregrounding, some of which have received experimental support. A broad classification of these influences would divide them into factors which operate at the point at which the antecedent is introduced, and factors which operate within the following text. Thus, an entity may be immediately signalled as important because of its occurrence in sentence subject position, or because of its level in the text hierarchy; or alternatively, its

importance may only become evident in the following sentences, as it continues to be mentioned or to play a large part in the action. These two types of influence are presumably independently variable, and the manner in which they combine together or interact to yield a total effect would be an interesting question. For present purposes, however, it is the latter type of influence which is most relevant. Distance itself is an influence occurring after the item has been mentioned, and it was felt desirable to manipulate foregrounding in a similar way; while in addition, such a manipulation had been successfully used by Lesgold, Roth and Curtis (1979) in their study of definite noun phrase anaphora. Details of the manipulation used here are given in the Method section, but in essence it consisted of writing intervening material which either was or was not connected to the antecedent, though in neither case mentioning that entity explicitly. Whether this manipulation would affect pronoun interpretation as it has been shown to affect definite noun phrase interpretation was itself a question of some interest.

As for Experiment 2, the present experiment assessed ease of pronoun interpretation by means of both the reading time technique and the pronoun question technique. Although in the last experiment the two measures produced corresponding results, the conclusion that they reflect the same processes is called into question by the subsequent discussion of the sources of the distance effect. That discussion suggested that two separate influences may have been at work, and the present experiment attempts to demonstrate this; but the possible operation of two distinct influences means that it cannot be determined which of them affected which dependent variable. The interesting possibility is that the two measures may show different patterns of effects when the factors assumed to be operating on them are carefully

separated. A discrepancy between the two measures might assist in deciding whether pronoun question accuracy reflects on-line processing difficulty or, as argued earlier, deferred inference processes.

The design of this experiment was somewhat simpler than that of Experiment 2, in that effects on pronoun interpretation were the only variables of direct interest; the two memory effects studied in the previous experiment were not investigated further. Those two tests had been included mainly as pilot work concerning possible sources of individual differences in reading ability, and as no evidence of any skill interaction was found they were no longer of interest. The reading span test was also discarded, as this had shown no connection with either reader skill or ease of pronoun interpretation. Another change from Experiment 2 was in the manner of assigning the experimental passages to each condition. In that experiment, each passage was written specifically for one condition, but in the present experiment the passages were allowed to vary so that they could each appear in all conditions. In this way, the danger of specific materials effects was reduced; as will be noted below, the fact that the intervening sentences necessarily vary in the different distance conditions means that the danger of specific effects is not wholly removed.

To summarise, Experiment 3 was concerned with three main questions. Firstly, antecedent foregrounding was manipulated in a manner similar to that used by Lesgold, et al. (1979), in an attempt to discover whether pronouns were sensitive to this variable as well as definite noun phrases. Secondly, foregrounding and distance were independently varied, in order to assess whether distance as such had any separate effect on ease of pronoun interpretation. Finally, the possibility that poor readers may be more sensitive than good readers to any effect of

distance, but not to an effect of foregrounding, was also considered.

Method

Subjects

24 members of the public, ranging in age from mid-teens to early forties, were recruited through a newspaper advertisement, and were paid for their participation. Subjects were assigned to skilled and less-skilled reader groups after the experiment was completed; as for Experiment 2, this division was done solely on the basis of their reading speeds.

Materials

Twenty experimental passages were written. Eight of these were loosely based on scenes and events from passages used in Experiment 2, while the remainder were completely new. The passages were written such that each one could take any of nine possible forms. An example passage is given in Table 3.1, and all twenty passages are given in full in Appendix 6.

Table 3.1

Experiment 3: Example Passage

The furniture store on Bridge Street caught fire early on Tuesday morning.

The blaze was discovered by the night watchman, who called the fire brigade

In minutes, the street was full of fire engines, police cars, and ambulances.

The fire was very dangerous, because of fumes from the artificial fibres.

An ambulanceman took care of the watchman, who had been affected by smoke.

A bystander offered to help, but a policeman told him to keep clear. The manager of the store, who had been notified, arrived in a little while.

A fireman began hacking at the side door with an axe, but made little headway.

FOREGROUND BLOCK

The door was solid, thick, and made a formidable barrier.

The rain of heavy axe blows barely dented the hard wood.

Usually, it was easiest to chop away either the lock or the door hinges.

This time, neither part of the door seemed likely to give way.

Despite the heat and showers of sparks, the effort had to be kept up.

The door had to be broken down in order to get into the building.

As the axe swung against the door, it made a great thudding noise.

There was now a small splintered hole near the door handle.

BACKGROUND BLOCK

At the front of the building, flames were licking through the windows.

Houses on the other side of the street were having to be evacuated.

Small children, still asleep, were wrapped in blankets and carried away.

No-one had time to save any of their valuables or other possessions.

At each end of the street, barriers had been erected to keep back the crowds.

For some reason, a big fire will always attract lots of onlookers.

The flashing blue lights of police cars played across their faces.

Everyone was warmly dressed, because of the cold night air.

He stood back at last and shouted for a crowbar.

- 1) Who shouted for a crowbar?
- 2) Where was the store?
- 3) Why was the fire especially dangerous?

Each passage contained an initial section of several sentences, which was constant across all nine forms. The length of this initial section varied between passages. The last sentence of the initial section introduced the entity which was to be the crucial antecedent; these entities were people, animals or things, and were introduced either by a proper name or by a full noun phrase. They were mentioned only once, in the antecedent sentence; and the preceding sentences introduced several other entities of the same type. As in Experiment 2, this was done in order to make the eventual choice of pronoun referent a more demanding task. In the example passage, the crucial antecedent is the fireman trying to break into the burning building, while the alternatives include a policeman and a bystander. In some passages, the antecedent sentence also included mention of other plausible alternatives in addition to the crucial antecedent itself. The final sentence of each passage, which was also held constant across all possible forms, contained a pronoun which referred to the crucial antecedent entity; these pronouns were made unambiguous on semantic or pragmatic grounds, in the same way as was done in Experiment 2. In nineteen of these target sentences the pronoun was the first word of the sentence, while in one it came near the middle of the sentence. The pronouns used were 'he'(10 times), 'she'(5), 'it'(4), and 'his'(1).

The nine possible passage forms were constructed as follows. One of the possible forms of each passage was created by placing the final, target sentence immediately after the introductory section; this meant that there was no gap between the target and antecedent sentences. The other eight possible forms were constructed using additional, variable passage sections. For each passage, two additional blocks of eight sentences were written, either of which could be inserted, in whole or in part, between the target and antecedent sentences. One of these

blocks of eight sentences was written with the intention of keeping the crucial antecedent in the foreground, while the other was written with the intention of pushing the antecedent into the background. The Foregrounding factor was thus determined by which block of sentences intervened between the target and the antecedent. The blocks were written such that either the first 2, first 4, first 6, or all 8 sentences could be inserted, with the sentences always appearing in the same order. The Distance factor thus had levels of 0, 2, 4, 6, and 8.

The two blocks of intervening sentences differed as follows. In the antecedent sentence, the crucial entity was always characterised as being involved in a certain action or possessing a certain attribute. In each passage, the foregrounding block of sentences continued to relate to aspects of that particular action or attribute, while the backgrounding sentences moved the narrative away to new topics or locations. Thus, in the example passage, the foregrounding sentences describe such things as the thickness of the door which the fireman is attempting to break down, the reasons why it was essential to break it down, and the immediate circumstances of the activity. The backgrounding sentences, on the other hand, although staying within the general setting of the fire, direct the story elsewhere, by describing the evacuation of nearby streets, and the crowds looking on. Given the state of present understanding of the foregrounding phenomenon, it was clearly impossible to attempt to equate the degree or strength of foregrounding and backgrounding across passages; for this reason, no attempt was made, for example, to equalise the number of new topics introduced in the backgrounding sentences. It may also have been the case that the strength of the manipulation varied even through the blocks of sentences used in a single passage. For the present

experiment, it was considered sufficient to ensure a broad difference between the two blocks of sentences for each passage, in order simply to establish the existence of the general foregrounding effect. Refining the details of how this effect came about would then be a problem for future research.

As shown in Table 3.1, the passages were followed by three questions. The first of these interrogated the referent of the pronoun in the target sentence, in the manner described for Experiment 2, while the second and third questions interrogated other information in the passage. No attempt was made to deliberately vary whether these latter questions interrogated facts or inferences. Given that the central portion of each passage was variable, the second and third questions always interrogated information from the constant introductory section. The questions were thus always the same, whichever form a passage took on a particular presentation. In one passage a question addressing information only mentioned in the backgrounding block of sentences was unintentionally included; this, however, was an isolated occurrence, and subjects who saw this passage in a form in which the required information had not in fact been presented generally produced answers consistent with the overall passage content. In order to conceal the patterns in the experimental passages, ten distractor passages were used; these were the same passages as were used for that purpose in Experiment 2. Also taken from Experiment 2 were the fourteen concealed practice passages.

From the twenty experimental, ten distractor, and fourteen practice passages, two random passage sequences were constructed for each subject. Firstly, for each subject two experimental passages were randomly assigned to the eight conditions given by the crossing of the

Foregrounding and Distance factors, while four passages were assigned to the ninth, zero-gap condition. These randomisations were done separately for each subject, with no attempt to balance the number of times each passage appeared in each condition. In the zero gap condition, the Foregrounding factor as manipulated in this experiment was clearly not applicable; however, in order to simplify the experimental design and analysis, it was introduced as a dummy factor, by randomly assigning two of the zero-gap passages to the foregrounded condition and two to the backgrounded condition. This allowed the data to be analysed as a complete $2 \times 2 \times 5$ design, as will be described below, with two skill groups and ten experimental conditions.

When two passages had been randomly assigned to the ten conditions for each subject, one passage from each condition was randomly chosen for inclusion in the first of the two passage sequences to be seen by that subject. The other experimental passages were included in the second passage sequence. For each subject, the first passage sequence read always began with the seven concealed practice passages used in the first sequence of Experiment 2; similarly, the second passage sequence read by each subject always began with the seven concealed practice passages used in the second sequence of Experiment 2. The concealed practice passages were always presented in the same order, but were followed by different random orderings of five distractor passages and the ten experimental passages assigned to any particular sequence. The first sequence read by each subject always included the five distractors used in the first sequence of Experiment 2, while the second sequence read by each subject always included the five distractors used in the second sequence of Experiment 2. There were thus two distinct twenty-two passage sequences for each subject, with the same practice and distractor passages used for each subject but each containing a

different random assignment of experimental passages to conditions. The materials were randomised in this way in order that the subjects could be divided into skill groups after the experiment was completed; if the materials sets had been balanced, it would have been necessary to pre-select the skilled and less-skilled readers in order to ensure that the two groups were presented with comparable experimental materials sets.

As the final aspect of the materials, the two passages used in Experiment 2 as measures of reading speed were again used for that purpose in this experiment. As for the distractor and practice passages, the reading speed passage which had been assigned to the first sequence of Experiment 2 always appeared in the first sequence in the present experiment, with the other reading speed passage always appearing in the second sequence.

Procedure

As for Experiment 2, the experiment was conducted over two separate one-hour sessions, with subjects reading one sequence in each. There was normally a gap of one hour between sessions, during which subjects carried out a completely unrelated experiment, but some subjects did go through the two sessions with only a five-minute break.

The sequences were presented in exactly the manner described for Experiment 2. The only difference was the omission of the sentence recognition test and the reading span test. As before, in each sequence the reading speed passage was interpolated after the fourth concealed practice passage.

Results

Subjects were assigned to skilled and less-skilled reader groups on the basis of their mean full-screen reading speeds, as described for Experiment 2. In the present experiment, times on the two passages had a product-moment correlation of +0.585; this was somewhat lower than the correlation obtained in Experiment 2, but was still highly significant ($df=22$, $p<0.005$).

Main Analyses

The target sentence reading times were analysed as a $2 \times 2 \times 5$ design, with Skill as a between-subjects factor and Distance and Foregrounding as within-subjects factors. As each subject read two passages in each condition, the analysis was conducted on the means of these two values. The ANOVA summary table for this analysis is given in Appendix 16. There was a significant main effect of Skill, with $F(1,22)=7.56$, $p<0.025$; mean target sentence reading time was 2771msec for the skilled reader group, and 3964msec for the less-skilled reader group, thus confirming the efficacy of full-screen reading speed as a predictor of sentence by sentence reading speed. There were, however, no interactions involving the Skill factor; for all three possible interactions, $F<1$. Table 3.2 therefore presents the mean target sentence reading times for the ten experimental conditions taken across

both groups.

Table 3.2

Experiment 3: Mean Target Sentence Reading Times(msec)

	<u>DISTANCE</u>					
	<u>0</u>	<u>2</u>	<u>4</u>	<u>6</u>	<u>8</u>	
<u>FOREGROUND</u>	2597	2575	2953	3370	2966	2892
<u>BACKGROUND</u>	2517	3606	4498	3905	4689	3843
	2557	3090	3726	3637	3827	

There was a significant main effect of Distance, with $F(4,88)=6.24$, $p<0.001$. From Table 3.2 it can be seen that reading times increased from 0-2, and again from 2-4, but remained fairly constant at 4, 6, and 8. Newman-Keuls tests showed that in fact neither the 0-2 nor the 2-4 increases were significant; however, the 0-4 comparison was significant, with $Q(4)=5.46$, $p<0.01$. Distance 0 was also significantly different from 6 ($Q(3)=5.04$, $p<0.01$), and from 8 ($Q(5)=5.93$, $p<0.01$); there were no significant differences between 2, 4, 6, and 8 (all $df=88$). It thus appears that increasing target-antecedent distance led to an increase in reading times up to an eventual asymptotic level. However, the nature and sources of the effect of Distance are made clearer by considering its relation to the effect of Foregrounding.

Firstly, there was a strong main effect of Foregrounding, with $F(1,22)=31.40$, $p<0.001$. Target sentences were read almost a second faster when the intervening material preserved the antecedent in the foreground than when it pushed the antecedent into the background. It

can be seen that the two conditions did not differ at distance 0, which was to be expected since the Foregrounding factor at that distance was in fact purely notional, but that at all other distances reading times were longer in the backgrounded condition. The simple main effects of Foregrounding were tested separately at the five levels of Distance. At 0, $F < 1$; at 2, $F(1,110)=5.23$, $p < 0.025$; at 4, $F(1,110)=11.55$, $p < 0.001$; at 6, $F(1,110)=1.46$; and at 8, $F(1,110)=14.53$, $p < 0.001$. The effect thus fails to appear only at distance 6, but returns even more strongly at distance 8.

In addition to the main effect of Foregrounding, there was a significant Distance \times Foregrounding interaction, with $F(4,88)=2.48$, $p < 0.05$; it should be pointed out, however, that this important result would not be considered significant if conservative degrees of freedom were used. It can be seen that while there was an increase in reading times of over a second from 0-2 in the backgrounded condition, in the foregrounded condition there was essentially no difference between distances 0 and 2. It may also be noted that from 4-6 there was an increase in reading times in the foregrounded condition, but a decrease in reading times in the backgrounded condition. Neither of these latter changes was significant when tested separately, but the combination of these two random fluctuations may explain the failure to obtain a significant Foregrounding effect at distance 6.

The simple main effect of Distance was not significant in the foregrounded condition ($F(4,176)=1.07$), but was significant in the backgrounded condition, with $F(4,176)=7.33$, $p < 0.001$. Newman-Keuls tests were therefore conducted between the five levels of Distance in the backgrounded condition, and yielded a pattern of significant differences similar to that obtained on the overall means. The 0-2 increase was not

significant, and nor was the 2-4 increase. For the 0-4 comparison, $Q(4)=4.41$, $p<0.05$, while for the 0-8 comparison $Q(5)=4.83$, $p<0.01$. The 0-6 comparison was not significant, and there were no significant differences between 2, 4, 6, and 8.

It thus appears that the effect of Distance was due almost entirely to its effect in the backgrounded condition. Reading times were not significantly affected by the number of intervening sentences, if these sentences held the antecedent in the foreground; but reading times did increase as distance increased, if the intervening sentences pushed the antecedent into the background. Such an interaction exactly fits the predictions of the working memory theory of foregrounding; topical continuity through the intervening sentences acted to maintain undiminished antecedent accessibility to pronominal reference, while topical discontinuity in the backgrounded case led to an immediate fall in accessibility. It may also be noted that the effect of Distance was due solely to the comparisons with the zero gap condition; when the data were analysed as a 2 x 4 design, with the zero gap condition omitted, both the main effect of Distance and the Distance x Foregrounding interaction proved nonsignificant. This confirms the finding that there were no significant pairwise comparisons between 2, 4, 6, and 8. Thus, additional distance produced no effect on ease of pronoun interpretation.

One possible difficulty with the analysis described above concerns the appropriacy of the random assignment of passages at distance 0 to the foregrounded and backgrounded conditions. This allowed the 2 x 5 analysis to be carried out, and in fact the two mean reading times at distance 0 did prove to be approximately equal, as expected; however, it could be argued that any random difference in the two reading times

at that distance would have had a distorting effect on the analysis, and in particular on the nature of the Distance x Foregrounding interaction. Although the danger of this did not seem great on the present occasion, it was felt desirable to reanalyse the reading time data in a manner which would allow a single mean at distance 0 to be directly compared with the other eight conditions; in this way, the validity of the conclusion that backgrounded conditions differed significantly from the zero-gap condition, while foregrounded conditions did not, could be reassessed.

To this end, a single mean reading time at distance 0 was calculated for each subject, taken across all four passages read at that distance. Similarly, a single mean was calculated for each of the other eight genuine conditions, taken across the two passages read in each condition by each subject. These nine mean values were then analysed as a 2 x 9 design, with Skill and Condition as factors, and the ANOVA summary table is given in Appendix 16; the overall means for the different conditions were of course identical to those given in Table 3.2. The analysis again revealed the significant main effect of Skill, with $F(1,22)=7.49$, $p<0.025$, while the main effect of Condition also proved significant, with $F(8,176)=5.92$, $p<0.025$; the interaction was not significant, with $F<1$. Given the significant effect of Condition, Newman-Keuls tests were used to compare the mean at distance 0 with the other eight means. It was found that none of the foregrounded conditions differed significantly from zero. In the case of the backgrounded conditions, for the 0-8 comparison, $Q(9)=6.62$, $p<0.01$; for the 0-4 comparison, $Q(8)=6.03$, $p<0.01$; and for the 0-6 comparison, $Q(7)=4.19$. The latter comparison was only marginally significant, while the 0-2 comparison was clearly nonsignificant ($df=176$ for all comparisons).

These results have therefore generally confirmed the findings of the earlier analysis. Distance 0, even when estimated over all four available reading times rather than over a randomly selected pair of reading times, was still not significantly different from any of the other distances in the foregrounded case. In addition, the pattern of significant differences between distance 0 and the four backgrounded conditions was repeated; only the 0-4 and 0-8 comparisons proved significant, with the others being marginal at best. The original procedure, with distance 0 randomly divided into foregrounded and backgrounded conditions, is therefore supported.

The analysis of the pronoun question data will now be reported. The ANOVA summary table for this analysis is contained in Appendix 16. Table 3.3 presents the mean number of correct answers (maximum=2) to the pronoun questions in the ten experimental conditions; there was no main effect of Skill, and again no interactions involving that factor (all $F < 1$). There was, however, a main effect of Distance, with $F(4,88)=2.84$, $p < 0.05$; accuracy on the pronoun questions fell, as the target-antecedent distance increased. Newman-Keuls tests showed that only the 0-6 and 0-8 comparisons were significant, with $Q(4)=3.84$, $p < 0.05$ in each case (both $df=88$). This effect was thus not a complete reflection of the effect of Distance on target sentence reading times, yet the correspondence was reasonably close. It is noteworthy especially that on both measures the effect seems to reach an asymptotic level. However, the distance effect on pronoun question accuracy would fail to reach significance with conservative degrees of freedom.

The correspondence between the reading time and accuracy results failed to appear in relation to the Foregrounding factor. There was no main effect of Foregrounding on question accuracy, with $F < 1$; and nor

Table 3.3

Experiment 3: Mean Pronoun Question Accuracy (max=2)

	<u>DISTANCE</u>					
	<u>0</u>	<u>2</u>	<u>4</u>	<u>6</u>	<u>8</u>	
<u>FOREGROUND</u>	1.792	1.667	1.708	1.458	1.458	1.633
<u>BACKGROUND</u>	1.750	1.708	1.583	1.542	1.458	1.608
	1.771	1.687	1.646	1.500	1.500	

was the Distance x Foregrounding interaction significant, also with $F < 1$. The failure to obtain an effect of Foregrounding on pronoun question accuracy goes against the prediction that both measures would reflect any change in the ease of pronoun interpretation.

Subsidiary Analyses

The above results were those concerned directly with the main questions of the experiment; however, various other features of the experiment were analysed, and these will now be reported.

Firstly, comprehension performance on the second and third questions attached to each experimental passage was analysed. Scores on these two questions were combined, and analysed as a $2 \times 2 \times 5$ design, with Skill, Foregrounding and Distance as factors. It was not expected that experimental condition should affect comprehension performance, and this was found to be the case; however, skilled readers were found to be significantly more accurate than less-skilled readers, with $F(1,22)=5.42$, $p<0.05$. For skilled readers, mean accuracy (maximum=4) was 3.300, while for less-skilled readers it was 2.942. The skilled readers thus not only read faster, but also performed more accurately on the comprehension questions.

The second additional analysis concerned the reading times for the intervening foregrounding and backgrounding sentences themselves. Although the finding that pronoun sentences were read faster after intervening foregrounding material was predicted on theoretical grounds regarding the construction of a pronoun-antecedent link, McKoon and Ratcliff (1980) make a criticism of the reading time technique which casts doubt on this interpretation. They point out that changes in target sentence reading times may be due to 'spillover' effects from the changed context; longer reading times for preceding context sentences may lead to longer reading times for the target sentence, due to continuing processing of the earlier sentences. On this view, the

longer target sentence reading times in the backgrounded condition may have been caused by longer reading times for the backgrounding sentences themselves.

To test this possibility, a comparison was made between subjects' reading times for the two types of intervening material. For each subject, a mean reading time was calculated for each of the eight intervening sentence positions, separately for the foregrounded and backgrounded conditions. In order that each mean should be based on the same number of raw reading times, only data from the passages presented at distance 8 were used; the sixteen data points for each subject were thus based on two reading times each. The analysis constituted a $2 \times 2 \times 8$ design, with Skill, Foregrounding and sentence Position as factors. It was found that there was a significant main effect of Skill, with $F(1,22)=6.82$, $p<0.025$; mean sentence reading speed was over a second faster for skilled readers than for less-skilled, but there were no significant interactions involving the Skill factor. The main effect of Position was not significant, with $F<1$, but the main effect of Foregrounding did prove significant, with $F(1,22)=4.79$, $p<0.05$. Mean sentence reading time for intervening foregrounding material was 3148msec, while the corresponding mean for backgrounding material was 2807msec. Thus, contrary to the view that the obtained effect on target sentence reading times may have been due to longer reading times for the backgrounding sentences themselves, these results show that backgrounding sentences were in fact read over 300msec faster than foregrounding sentences. Whether the faster reading times for backgrounding sentences could have made some other contribution to the main experimental results will be returned to in the discussion section.

An interesting feature of the above analysis led to the carrying out of a second comparison between foregrounding and backgrounding sentences. Although the Foregrounding x Position interaction was not significant, with $F(7,154)=1.57$, it was noticeable that only at the first sentence position were foregrounding sentences read faster than backgrounding sentences, the reverse of the overall main effect. Sanford and Garrod (1981) report that sentences constituting a topic shift, such as sentences expressing a large time change, are read more slowly than sentences which do not indicate such a discontinuity. In the present experiment, the first backgrounding sentence constitutes a topic change, while the corresponding foregrounding sentence continues on topics related to the preceding antecedent sentence. It was considered possible that the reversal of the overall trend at the first sentence position may have been a reflection of this difference. To test this possibility more fully, a mean reading time was calculated for each subject for the first intervening sentence position, separately for foregrounding and backgrounding sentences; these means were calculated over all the passages which had been presented with intervening material, so that each of the two means was based on eight reading times. These mean values were analysed as a 2 x 2 design, with Skill and Foregrounding as factors. The usual effect of Skill was present, with $F(1,22)=8.33$, $p<0.01$, but there was no main effect of Foregrounding, with $F<1$. This analysis therefore failed to show any evidence that readers responded differently to the first foregrounding and backgrounding sentences.

The final analysis to be reported concerned the effect of sentence serial position within a passage. In Experiment 1, it was found that later sentences within a passage were read more slowly than the earlier sentences; by contrast, Experiment 2 showed that subjects read more

quickly as they progressed through a passage. The analysis in the present experiment was a complete replication of that carried out in Experiment 2, using the same design and taking the reading time data from the same concealed practice and distractor passages. The outcome was virtually identical to that obtained in the previous experiment. There was a significant main effect of sentence serial position, with $F(16,352)=17.06$, $p<0.001$; as for Experiment 2, the most striking aspect of the pattern of reading times was the very long mean reading time for the first serial position, of over six seconds. Mean reading time for the second sentence was only just over four seconds, with reading times then declining more gradually up to the seventh serial position. Times then stabilised at approximately three seconds. As in Experiment 2, readers spent longest on the first sentence of a passage, and then increased in speed to an eventual asymptotic level.

Discussion

This experiment has clearly shown that pronoun interpretation is sensitive to the foregrounding effects of topical continuity. Pronoun sentence reading times were much faster when the material separating them from their antecedents dealt with topics closely related to that antecedent, than when the intervening sentences dealt with new and unrelated topics. In line with the working memory theory, it can be argued that topical continuity led to antecedents being retained in some more active and immediately accessible state, owing to their continued importance and relevance to the text. In the foregrounded condition, processing of the intervening sentences presumably involved making repeated connections to the crucial entity itself, while in the

backgrounded condition the sentences could be integrated without involving that specific entity. Items may therefore be foregrounded because of their implicit involvement in current text processing, rather than simply being maintained because the text suggests that an explicit reference to them may shortly occur.

This experiment also revealed an overall effect of pronoun-antecedent distance on reading times. This effect was more regular than that found in Experiment 2, consisting as it did of a rise in reading times to an eventually stable asymptote; however, the most important conclusions regarding the distance effect are based on the interaction of distance and foregrounding. Reading times were not significantly affected by distance in the foregrounded condition; provided the antecedent continued to be relevant to the text, there was no loss in availability to pronominal reference even after eight intervening sentences. In the backgrounded condition, reading times showed an immediate increase of almost a full second at distance 2, but beyond this point showed no clear trend towards further increases; although reading times fluctuated, the longer distances showed no significant differences from distance 2. The lack of distance effect in foregrounded conditions can be attributed to continued working memory presence of the crucial antecedents. Since additional distance did not lead to additional reading time increases in the backgrounded condition, and since distance had no effect whatever in the foregrounded condition, it may be concluded that distance as such has no direct influence on pronoun interpretation; obtained distance effects may be attributed to confounded topical changes. Antecedent availability is governed by topical relevance, and appears to vary in an all-or-none fashion; it shows no gradual effect of surface distance.

There are two possible difficulties with these conclusions. One of these is the danger, alluded to in the Method section, that the strength of the foregrounding or backgrounding manipulation may not have been constant through the blocks of intervening material. As Lesgold, Roth and Curtis (1979) showed, it is quite possible for an antecedent to become foregrounded again even without being explicitly mentioned; inclusion of two sentences referring to concepts linked to the antecedent, at the end of a block of four backgrounding sentences, was found there to lead to faster target reading times. In the present experiment, the possibility exists that the later sentences in the intervening blocks may have been more closely related to the antecedent than the earlier sentences; if this were so, any effect of distance may have been obscured. One way of testing this possibility would have been to carry out a ratings test similar to that reported in connection with Experiment 2, assessing the salience of each antecedent at each point in the intervening blocks. However, it was felt that with twenty passages in the experiment, all of which appeared in all conditions, the probability of such random variations in foregrounding accounting for the results was extremely slight. Clearly, to establish the validity of the results a replication on new materials would be desirable; Experiment 6 in fact presents such a replication.

The second difficulty concerns the differences in reading time for the intervening foregrounding and backgrounding sentences themselves. Reading times for these sentences were analysed in order to establish whether longer reading times for backgrounding material could have been a possible cause of longer target sentence reading times, by means of some form of spillover effect; in fact, it was found that backgrounding sentences were read over 300msec faster than foregrounding sentences. However, a spillover effect is still a possibility. Pronoun

interpretation involves not only locating an antecedent, but also integrating the new information with the representation of the previous text; if the longer reading times for the foregrounding sentences reflect the construction of a more detailed and better organised representation, then target sentence integration may have been speeded independently of effects on antecedent location. This possibility relates to the notion of scenario richness discussed by Sanford and Garrod (1980; also Garrod and Sanford, 1981). While this suggestion is purely speculative, it offers a possible alternative explanation for the main effect of foregrounding.

Both the above difficulties will be addressed by replications of the foregrounding effect to be described in later experiments. For the moment, the present reading time results may be summarised as indicating that topical discontinuities in a text produce an all-or-none decrease in antecedent availability, and that apparent effects of surface distance are due to confounding with such discontinuities. These are the main conclusions of interest here, but the various other aspects of the experiment will now be discussed.

A striking feature of the experiment was the lack of correspondence between the reading time results and the pronoun question results. Pronoun question accuracy did show a significant main effect of distance, which was reasonably similar to that obtained on reading times; however, question accuracy showed no main effect of foregrounding, and nor did distance and foregrounding interact. These differences in sensitivity to the two independent variables run counter to the view that the two measures reflect the same psychological processes. It seems reasonable to argue that reading time should be an accurate measure of ease of pronoun interpretation, since it directly

reflects the difficulty of sentence processing; however, if it is assumed that the pronouns are in fact eventually assigned to their antecedents, there is no obvious reason why question accuracy should be affected by the length of time taken to achieve this assignment. This argument could be put forward to explain the lack of a foregrounding effect on accuracy, but it does not explain why accuracy shows an effect of distance. A more adequate explanation is that already noted in the earlier discussion of the paper by Daneman and Carpenter (1980); pronoun question accuracy may be affected by inferences made at the time of answering, as well as by whether or not the pronoun was successfully assigned. To take again the example from Daneman and Carpenter, a question such as "Who finally arrived?" can be answered by remembering which individual was described as being late for a meeting; answers arrived at in this way need make no use of the pronoun sentence itself, but only of the information contained in the antecedent sentence. Since such inferences are not part of discourse comprehension processes, but involve explicit use of the memory representation after the text has been read, it is not contradictory to argue that they are made more difficult as a result of the location in the passage of the required information. In Experiment 2, it was found that pronoun questions were answered less accurately if the target sentences themselves were further back in the passage; here, it is argued that it is the recency of antecedent information which is in fact crucial. Thus, while the pattern of reading time results shows sensitivity to foregrounding, the accuracy results were sensitive only to distance; and this is explicable on the view that reading times reflect assignment processes, while accuracy reflects deferred inference processes.

In addition to considering the effects of distance and foregrounding themselves, and to comparing the two measurement

techniques, this experiment was also concerned with the possibility that reader skill may affect sensitivity to the distance effect, but not to foregrounding. Since it has been shown in the above discussion that distance as such in fact has no effect on pronoun interpretation, it is perhaps unsurprising that distance and skill showed no interaction. However, the absence of a skill interaction with foregrounding is unexpected, since the research reviewed earlier suggests that poor readers are less able to identify discourse topics than good readers.

The present set of experiments has produced a number of results which can be interpreted within a working memory theory of pronoun interpretation. This class of theories is represented by models such as those of Kintsch and van Dijk (1978) and Sanford and Garrod (1981), and argues that pronoun assignment involves a first processing stage consisting of an attempted direct match to some form of immediately accessible store. Pronoun interpretation will be easy and rapid if the intended antecedent is available in that store, but will be difficult and time-consuming if it is not. Various influences have been claimed to affect whether or not an antecedent will be immediately accessible at any given point in a text; the present experiments have considered the contributions made by pronoun-antecedent distance and by topical continuity or foregrounding.

That pronoun-antecedent distance as such was found to have no influence on ease of pronoun interpretation is a reasonable outcome. Two mechanisms have been proposed to account for obtained distance effects; one of these, attributing the effect to the introduction of new topics and a decline in antecedent salience, is fully compatible with the results obtained here. The other mechanism attributed the distance effect to more general aspects of memory functions, such as decay and displacement, and would predict a distance effect independent of foregrounding; this mechanism must be rejected on the basis of the present results. Working memory contents in reading appear to be under constant strategic control, with items being selected for inclusion on the basis of their current topical relevance. In this way, any working memory system put forward as an explanation for these results must

clearly be highly flexible, and rapidly responsive to processing decisions.

Foregrounding was manipulated in Experiment 3 by means of topical continuity, in a manner similar to that used by Lesgold, Roth and Curtis (1979); sentences following the antecedent dealt with topics closely related to that entity, and it was suggested here that the antecedent was therefore implicitly involved in integrating these sentences. In this way, the degree of activation of the antecedent may depend on it being repeatedly used in processing of subsequent text. A similar mechanism may account for the height effect found by Cirilo (1981). Entities with a high level in the text hierarchy were there found to be easier to refer to than those at a relatively low level, even after three intervening sentences; presumably, higher-level entities are more likely to have been involved in integrating the subsequent text. However, it is not clear that all the reported foregrounding phenomena can be accounted for in this way. In particular, the findings reported by Sanford and Garrod (1981), that a large stated time change can produce an immediate fall in antecedent availability, is difficult to explain by this mechanism. Scenario-dependent entities were there shown to be more difficult to refer to after a sentence indicating that the action had moved outside the likely time range of that scenario, while reference to more globally topical entities was not affected. This seems to require some form of selective mechanism which makes active decisions on working memory presence, rather than the more passive mechanism described above. One question which may be asked is whether all forms of foregrounding phenomena occur by the same mechanism; it is possible that not all manipulations have similar processing consequences, even though the effects on target sentence reading time may be comparable. Rather than pursue this question here, the following

series of experiments will continue to deal with the foregrounding effect demonstrated in Experiment 3; although attempts will be made to link this to a more general account of foregrounding, no empirical evidence on this generality will be presented.

It should be pointed out that antecedent foregrounding offers one possible explanation for the failure to obtain a significant distance effect in Experiment 1. In that experiment, the crucial antecedents were the only specific individuals mentioned in each passage, and hence may have been given global topical status; in addition, the intervening filler sentences were fairly neutral in content, and so may not have constituted a clear topic shift. For both these reasons, the likelihood that the antecedent may have been actively retained in working memory throughout the text must have been increased.

The lack of interactions between reader skill and either distance or foregrounding is puzzling. An interaction between skill and foregrounding in particular would be strongly predicted on the basis of the results reviewed earlier showing poor readers to make less use of text structure and to be less well able to identify discourse topics; on such a basis, poor readers would be expected to be less sensitive to foregrounding than good readers. However, the absence of such an effect in the present research should not be seen as conclusive. It is possible that the present experiments were simply insufficiently sensitive to skill differences owing to the nature of the ability divisions which were employed. As noted in Experiment 1, the method of dividing subjects at the median of some skill measure does not guarantee a wide difference between the two groups, while in addition the present experiments used adult subjects of a presumably narrow range of ability. Even though subjects in Experiment 3 were recruited from the general

public, and therefore presumably contained a wider range of reading skills than a sample of student subjects, these skill differences are likely to have been smaller than those present in published studies using children. The lack of skill effects may therefore be due to sampling problems rather than to genuine absence of such effects in a wide enough population. For present purposes, reader skill is a subsidiary topic and the experimental improvements which might make skill interactions more likely would not be justified in the following investigation of foregrounding; nevertheless, reader skill will continue to be included as a factor in the next set of experiments.

Some form of working memory theory would therefore appear to offer a reasonable interpretation of the results of these three experiments. However, there are many details of such a theory of pronoun assignment which remain to be worked out, and a number of important problems left unsolved. One obvious issue concerns the nature of the short-term storage system which is involved; various suggestions have been made, and although these may have different processing implications if made sufficiently precise there is as yet no basis on which to choose between them. Similarly, the nature of the information stored, and its manner of representation, are unspecified; whether working memory holds a set of complete propositions, as suggested by Kintsch and van Dijk (1978), or a representation of nodes and pointers as suggested by Sanford and Garrod (1981), or some other form of information, remains to be determined. There are also many aspects of foregrounding phenomena which need to be investigated. One question concerns the maximum duration of the foregrounding effect found in Experiment 3; it seems unlikely that a text could continue indefinitely without mentioning the topical entity. Another issue concerns the different influences on antecedent salience, and the ways in which they interact; the outcome

of conflicting foregrounding signals, in which two or more entities are simultaneously marked as salient by different means, would give some indication of the relative strengths of the known influences.

These are merely a few of the areas in which a working memory theory of foregrounding could be extended and made more precise. The next two experiments, however, set out to make a test of the working memory theory of foregrounding which does not rely on measures of ease of pronoun interpretation.

Chapter III

Recognition Memory Tests of Foregrounding

- 1 Experiment 4
- 2 Experiment 5
- 3 General Discussion

Introduction

This experiment set out to make a direct test of one crucial aspect of the working memory theory of foregrounding. That theory argues that foregrounded antecedents are maintained in a directly accessible store throughout the intervening material, with the eventual pronoun being easier to interpret because the required referent is already activated. Pronoun sentence reading time is thus an indirect reflection of antecedent activation, or presence in working memory. A more direct method of assessing an item's degree of activation is by means of an explicit recognition memory test. If subjects are presented with a test word immediately after reading a passage, and asked to decide whether or not that word did in fact occur in the passage, performance on genuinely old items would be expected to be improved if the item was already activated rather than being present only in long-term memory. By this reasoning, it could be predicted that recognition tests to antecedent nouns should produce faster latencies and greater accuracy if the test is made at a point in the text where the antecedent is foregrounded rather than at a point where it is backgrounded.

A number of studies have made use of the recognition memory technique when investigating the activation of particular concepts during comprehension. Caplan (1972) used auditory presentation of simple sentences, followed by a single probe word for subjects to identify as either old or new; latencies to old words were faster if they had occurred in the last clause of the presented sentence rather than the first. This result corroborated other findings indicating that

information from the most recent clause is more available than that from earlier clauses (e.g. Jarvella, 1971). Chang (1980) extended the recognition technique to reading, in an explicit attempt to validate the method as one which could be reliably used to study active memory processes. His subjects read simple two-clause sentences on a computer screen, with the sentences being presented one word at a time; each sentence was followed by a single-word recognition test. In the first experiment, old words were again found to be recognised significantly faster, and with fewer errors, if they had occurred in the last clause rather than the first; in this way, recognition tests were shown to be sensitive to activation levels in reading. Chang's second experiment was concerned with pronominal reference. In each test sentence, the second clause contained either a noun repetition or a pronoun, referring back to one of two entities mentioned in the first clause; the crucial recognition test words were the original antecedent nouns, with the other of the two first-clause nouns acting as a control. It was found that recognition latencies were faster if the test word had been referred to again in the second clause than if it had not; pronouns produced less facilitation than complete noun repetition, but were still significantly faster than the control condition. It was concluded that use of the pronoun in the second clause had reactivated the meaning representation of its referent.

McKoon and Ratcliff (1980) have produced a similar result. They presented subjects with short passages in which the final sentence either did or did not refer back to an individual mentioned in the first sentence; the anaphoric terms in the final sentence were either repetitions of the original noun phrase, or superordinate category terms. After each passage, the antecedent noun was presented for recognition. As with Chang's experiment, latencies were faster if the

test word was referred to in the final sentence than if it was not, even when the anaphor was lexically quite distinct. In their second experiment, McKoon and Ratcliff investigated a question of perhaps closer relevance to foregrounding. Rather than being the antecedent nouns themselves, the recognition test words on this occasion were nouns which had occurred in the same proposition as the antecedent in the first passage sentence; it was found that recognition latencies to these linked nouns were also faster if the final passage sentence referred back to the antecedent. Thus, anaphoric reference to an entity resulted in greater activation not only of that entity itself, but also of concepts linked to it earlier in the text. This finding may have some bearing on the foregrounding manipulation in Experiment 3, since the intervening sentences there either did or did not continue to refer to topics linked to the crucial antecedent; on McKoon and Ratcliff's results, it could be argued that this reference to linked concepts should have led to the antecedents themselves also being activated, as the working memory theory proposes.

Another paper which has used the recognition technique in a way which may shed light on foregrounding processes is that of Fletcher (1981). On the basis of the Kintsch model of reading (Kintsch and van Dijk, 1978; Kintsch and Vipond, 1979), Fletcher made predictions as to which words from a text would be in the short-term memory buffer at the end of each processing cycle; at various points in the text, subjects were presented with single words for recognition. A number of interesting results were found. Firstly, recognition latencies were fastest of all for test words which had occurred in the cycle immediately preceding the test itself; this was taken to indicate the availability of surface information at that point. Secondly, when the test words had occurred in the next to last processing cycle, latencies

were faster for those words which were predicted to be in the buffer than for those which were not; thus, on the basis of recency and text level involved in the leading edge selection strategy as discussed earlier, the Kintsch model successfully predicted the level of activation of particular items. Finally, it was discovered that latencies were equally slow for items which the model predicted not to be in the buffer, irrespective of whether those items had occurred in the next to last cycle or an earlier cycle. This result appears to be comparable to the absence of an additional distance effect on pronoun sentence reading times in backgrounded conditions.

There are thus a number of studies demonstrating the sensitivity of the recognition technique to the level of activation of text items, and some of these results bear on foregrounding. The results of McKoon and Ratcliff (1980) showed that items may be activated, even without being explicitly mentioned, if items linked to them earlier in the text are mentioned; while Fletcher (1981) showed that items selected on a criterion including consideration of text level are also activated. However, neither of these experiments tested recognition latencies for items which were known to be either foregrounded or backgrounded; there is no guarantee that the items found to have rapid recognition latencies would also have been easy to refer to pronominally. While these results are therefore indicative, an explicit test of recognition memory is required on materials known to produce a foregrounding effect. The present experiment carried out such a test, by using the passages from Experiment 3 and giving subjects recognition tests on the crucial antecedent nouns. If the working memory theory of foregrounding is correct, such recognition tests should be sensitive to the foregrounding manipulation already demonstrated to affect pronoun sentence reading times. The working memory theory thus predicts faster latencies and

greater accuracy on test words which are currently foregrounded.

Method

In most aspects of materials, design and procedure this experiment was identical to Experiment 3. The description below therefore only goes into detail on those features which were changed in order to investigate recognition test responses rather than pronoun sentence reading times.

Subjects

Subjects for this experiment were twenty-four first-year psychology students at Plymouth Polytechnic, who received course credit for their participation. They were divided into skilled and less-skilled reader groups after the experiment was completed, in the manner previously described.

Materials

The same twenty experimental passages used in Experiment 3 were used again in this experiment, with some slight modifications to take account of the changed methodology. These alterations were:-

(1) In four of the Experiment 3 passages (passages 5, 6, 7 and 15 -; these numbers identify the passages given in full in Appendix 6) the crucial antecedent noun phrase, subsequently referred to by the pronoun

target sentence, consisted of two words. It was felt desirable that all recognition tests in this experiment should be for single words, and therefore these four antecedents were changed. In passage 5, "chief engineer" was altered to "engineer"; in passage 6, "managing director" was altered to "chairman" (and the earlier reference to "chairman" was altered to "managing director"); in passage 7, "Bengal tiger" was altered to "tiger"; and in passage 15, "red setter" was altered to "labrador". In this way, all the crucial noun phrases were made into single words, either common nouns or proper names.

(2) In some cases, the crucial antecedent noun in one experimental passage also occurred in another experimental passage, although not as crucial antecedent; this was especially the case with proper names, which were occasionally used more than once. In order to avoid any effect on recognition memory for the crucial test words of having encountered that noun in some other passage, all such repetitions were removed. Each test word thus occurred only once in the experimental materials, as the crucial antecedent noun in its own passage.

(3) As this experiment was not concerned with pronoun sentence reading times, but with immediate recognition memory for the antecedent noun at the point where the pronoun would have been encountered, the final target sentences used in Experiment 3 were removed from the passages. The expressions 'antecedent' and 'intervening material' will, however, continue to be used for the sake of continuity with other experiments. Since the passages no longer ended with a pronoun sentence, the questions interrogating the referent of the pronouns were also deleted. The remaining questions for each passage were retained. There were thus only two comprehension questions for each text in this experiment.

The experimental passages were otherwise unchanged. As for Experiment 3, the twenty passages were randomly assigned to the ten experimental conditions, with the antecedent being followed by either 0, 2, 4, 6 or 8 sentences of intervening material, drawn from either the foregrounding or backgrounding block. As before, a separate randomisation was carried out for each subject, and the Foregrounding factor at distance 0 was a randomly assigned dummy variable. For each subject, one of the two passages assigned to each condition was randomly assigned to each of the two passage sequences that were to be presented.

The concealed practice and distractor passages used in Experiment 3 were also used again in this experiment. Those which had appeared in the first sequence of Experiment 3 were assigned to the first sequence of this experiment, while those which had appeared in the second sequence of Experiment 3 were assigned to the second sequence of this experiment. Since all the test words used with the twenty experimental passages were necessarily Olds, which had actually occurred in the passage, the main purpose of the distractor passages in this experiment was to provide test words that would have to be recognised as News, that had not in fact appeared in the passage. However, as some of the distractor tests would also have to be Olds, in order to avoid any possible pattern that might highlight the difference between the experimental and distractor passages, it was felt that the previous mix of five distractors and ten experimental passages would not be sufficient for this experiment. For this reason, four of the concealed practice passages in each sequence were randomly selected to be used instead as distractors in that sequence. In this way, each sequence contained only three passages for practice, but nine which could be mixed in with the experimental texts. It should be noted that this

reduction in the amount of reading practice was justifiable, since in this experiment sentence by sentence reading times were not of direct interest.

For each of the concealed practice and distractor passages, a single word was chosen to be presented as a recognition test. As indicated above, some of these test words were Olds, which were either common nouns or proper names that had actually occurred in the passage, while some were News. Of the New test words, some were plausible, in that they were to some extent related to the content of the passage, while some were noncase, in that they were relatively unconnected to the passage content. In both sequences, two of the concealed practice passages were tested with Olds, and one with a plausible New; for the distractor passages, in the first sequence three were Olds, four were plausible News and two were noncase, while in the second sequence there were two Olds, six plausible News, and one noncase. None of the practice or distractor passage test words were repeated anywhere in the materials, and nor did these passages contain repetitions of the experimental test words. Minor changes were made as appropriate to ensure that this was so. The test words chosen for these passages are given in Appendix 3. Finally, in order that all passages in this experiment should have only two comprehension questions, the third question on each of the practice and distractor passages was deleted.

The two passage sequences presented to each subject were constructed from these sets of materials in a similar fashion to that used in Experiment 3. Each sequence began with the three concealed practice passages, which were presented in the same order to each subject, and these were followed by a random arrangement of the nine distractors and ten experimental passages. This random arrangement was

different for each subject and each sequence, with the experimental passages being assigned to conditions as previously described. In each sequence, the reading speed passage, presented in the manner described for Experiment 3, was interpolated after the second concealed practice passage. These reading speed passages were the same as were used in Experiments 2 and 3, with the third question deleted and a recognition test word added; these test words are given in Appendix 4.

For each subject, there were thus two distinct passage sequences, with each containing one passage in each experimental condition. Finally, the two passages used for explicit practice in the previous two experiments were also used for that purpose here; they were given recognition test words, one Old and one plausible New, and these also can be found in Appendix 3.

Procedure

Each subject attended two one-hour sessions, in each of which one of the two passage sequences for that subject was read. The two sessions were attended either with only a five-minute break in between, or with a longer interval of up to a week.

The passages were presented sentence by sentence in the manner described for the previous two experiments. Instructions were given at the beginning of Session One. Subjects were told that it was an experiment on reading comprehension and memory, and that as well as comprehension questions on each passage there would also be memory tests for specific words. The format of the recognition tests was described, and subjects were told to decide whether or not the exact word presented

had occurred in the preceding passage. They were told to read the passages at their normal rate, and to make no extra attempt to memorise the exact wording of the sentences; and if they were unsure of the correct response on any particular test, they were instructed to make a guess. No mention was made of the importance of speed of responding on the recognition tests.

The format of the recognition tests was as follows. After the subject had pressed the red advance button after the last sentence of the passage, this sentence was cleared from the screen and there was a pause of half a second. A row of five asterisks was then displayed in the centre of the screen, and in half a second these were replaced by the test word. Test words thus appeared within one second of subjects completing reading the last sentence of the passage; the asterisks served as both a warning that the test was about to appear, and as a fixation guide indicating where the test word would be displayed. The test words were always displayed in lower case, although proper names had the usual initial capital letter. To give their decision, subjects pressed one of two buttons on the right of the response-key board in front of the screen; these were labelled 'OLD' and 'NEW' respectively. As soon as one of the buttons was pressed, the test word was cleared from the screen, and the two comprehension questions relating to that passage were displayed. Subjects wrote their answers to these questions in a booklet; recognition test responses and latencies were recorded automatically by the computer.

It should be noted here that a technical limitation on the computer system being used meant that no recognition latency of less than 578msec could be recorded. If the subject pressed a response key in a shorter time than this, the response could not be detected, and the subject

would have had to press the button a second time. No subject reported having to do this, however; in addition, the overall mean latency in the experiment turned out to be over two seconds, and therefore this limitation was presumably not a problem.

After being given the instructions at the beginning of the first session, subjects read through the two explicit practice passages in the presence of the experimenter, as training in the procedure. They were then left alone to read through the experimental sequence. At the start of the second session, subjects were simply told that it was identical to the first, but with different passages.

Results

Subjects were divided into reader skill groups on the basis of their reading speeds for the two passages which were presented on screen in full, paragraph format. The product-moment correlation between speeds on these two passages was $+0.782$ ($df=22$, $p<0.005$); this once again confirms the reliability of the reading speed measure, and the skill division was made on the mean of the two speeds. Those subjects with the twelve fastest mean reading speeds formed the skilled reader group, while those subjects with the twelve slowest mean reading speeds formed the less-skilled reader group.

Main Analyses

The recognition test data were analysed as a $2 \times 2 \times 5$ design, with Skill as a between-subjects factor and Foregrounding and Distance as

within-subjects factors; the recognition latency results will be reported first, followed by the accuracy results. In the analysis of the recognition latencies, only those latencies associated with correct responses were included; it should be noted, however, that reanalysis with all latencies included produced identical patterns of results. Since there were twenty-four subjects, each reading twenty experimental passages, a total of 480 recognition responses was recorded, of which only 24 were incorrect. The latency analyses were run on the GENSTAT statistical package which estimated values for these deleted data points. The ANOVA summary table given in Appendix 17 includes a factor labelled Trial, which represents the two passages read in each condition by each subject; this factor was included because means could not be calculated on this occasion, owing to the deletion of certain data points. Results involving the Trial factor were of no interest here.

Table 4.1

Experiment 4: Mean Recognition Test Latencies (msec)

	<u>DISTANCE</u>					
	<u>0</u>	<u>2</u>	<u>4</u>	<u>6</u>	<u>8</u>	
<u>FOREGROUND</u>	1649	2007	1884	2145	2327	2002
<u>BACKGROUND</u>	2445	2191	1933	1739	2636	2189
	2047	2099	1909	1942	2481	

The mean recognition latencies in the ten experimental conditions are given in Table 4.1. The main effect of Skill was found to be nonsignificant, with $F < 1$, and all interactions involving Skill also produced nonsignificant F values; the mean latencies relating to this

factor are therefore not presented. In Table 4.1, it can be seen that mean recognition latency in the foregrounded condition was 2002msec, while mean latency in the backgrounded condition was 2189msec; this difference in latency was in the predicted direction, but was found not to be significant, with $F(1,22)=2.69$. With regard to the effect of Distance, the most notable feature was that distance 8 produced a mean latency approximately 400msec slower than those obtained at any of the other four distances; however, the main effect did not even approach significance, with $F(4,88)=1.08$. All interaction effects produced F values less than or equal to 1.

It therefore appears that recognition latency was unaffected by either distance from the test word or the foregrounded status of that word. One possible problem with the analysis described above, however, concerns the mean latencies obtained at distance 0. As indicated earlier, the passages read at distance 0 were assigned to the foregrounded and backgrounded conditions on a purely random basis, and therefore these two means, as was in fact the case in Experiment 3, should not have greatly differed. On this occasion, however, the mean in the backgrounded condition was approximately 800msec greater than the mean in the foregrounded condition. This difference must presumably be due to chance, and has no interest in itself, but may possibly have had a distorting effect on the analysis. In particular, such a difference may have concealed any Distance x Foregrounding interaction such as was obtained in Experiment 3. For this reason, the latency data were reanalysed as a 2 x 9 analysis of variance, with Skill and Conditions as factors, in the manner described for the reanalysis of pronoun sentence reading times in Experiment 3. For each subject, the mean of the four observations at distance 0, and of the two observations in the remaining eight conditions, were used as data; in this way, a single mean for

distance 0 was directly compared with the eight means for the other genuine conditions. The resulting analysis, for which the summary table is given in Appendix 17, produced a nonsignificant main effect of Conditions, with $F(8,176)=1.75$. This confirms the lack of effects on the original analysis, and suggests that the random difference in latencies at distance 0 did not obscure any significant results.

The analysis of the recognition accuracy data will now be described; the ANOVA summary table is given in Appendix 17. As each subject read two passages in each condition, data points in the analysis could take the values 0 (both tests incorrectly responded to), 1 (one test incorrect), or 2 (both tests correct). As indicated earlier, there was a very high level of accuracy overall. In the analysis, it was found that the main effect of Skill was not significant, with $F=0$; the main effects of Foregrounding and of Distance were also not significant, with $F(1,22)=2.02$ and $F(4,88)=1.55$ respectively. The majority of the interaction effects produced F values of less than or equal to 1; there was, however, one significant exception. Table 4.2 presents the mean recognition test accuracy for the foregrounded and backgrounded conditions, and for the skilled and less-skilled reader groups; for the interaction of Skill x Foregrounding, it was found that $F(1,22)=5.61$, $p<0.05$. It can be seen that skilled readers were less accurate in the backgrounded condition than in the foregrounded condition, which is precisely the effect which was predicted overall; less-skilled readers, however, showed the reverse trend, being less accurate in the foregrounded condition than in the backgrounded condition. The simple main effects of the Foregrounding factor were tested separately for the two skill groups; it was found that for less-skilled readers, $F<1$, but for skilled readers, $F(1,44)=7.26$, $p<0.01$.

Table 4.2

Experiment 4: Mean Recognition Test Accuracy (max=2)

	<u>FOREGROUND</u>	<u>BACKGROUND</u>	
<u>SKILLED</u>	1.967	1.833	1.900
<u>LESS-SKILLED</u>	1.883	1.917	1.900
	1.925	1.875	

The Skill x Foregrounding interaction therefore appears to suggest that skilled readers were sensitive to the foregrounded status of test words, in the predicted manner, while less-skilled readers were not affected. However, there are two reasons for caution in evaluating this outcome, which are based on the nature of the data being analysed. Firstly, the narrow range of values taken by the data points may not be ideally appropriate for the analysis of variance technique, and secondly, the low frequency of errors, with only 5 per cent of responses being incorrect, could allow a purely random distribution to achieve significance. In addition, the interaction must be evaluated in the light of the preceding experiment, which showed no indication that skilled and less-skilled readers were differentially sensitive to foregrounding. The danger that this obtained interaction may be due solely to chance must therefore be considered.

In an attempt to evaluate this possibility, and to determine whether or not the interaction was based on a genuine skill difference, two additional analyses were conducted. Firstly, an analysis of variance similar to the original was carried out, using only the six fastest and the six slowest readers. In this way, the middle range of

subjects was eliminated, and thus any difference genuinely due to reader skill should have been emphasised. On this analysis, however, the Skill x Foregrounding interaction did not even approach significance, with $F < 1$. Secondly, a correlational test was calculated. For each subject, that subject's total number of correct responses in the backgrounded condition was subtracted from his total number of correct responses in the foregrounded condition; the resulting figure indicated the size and direction of any effect of the Foregrounding factor for that subject. This net effect of Foregrounding was then correlated with subjects' mean reading speeds on the two full-screen passages. The product-moment correlation was found to be -0.22, which was not significant. Thus two additional tests, one emphasising the extremes of the skill range and one making full use of all values along the continuum of reading speeds, failed to show any relation between the effects of Skill and of Foregrounding. Given this failure to find supporting evidence for the Skill x Foregrounding interaction which was obtained in the original analysis, and with the reasons for caution outlined in the previous paragraph, it may be concluded that the interaction did not in fact reflect a genuine skill difference.

Subsidiary Analyses

The above analyses, of recognition test latency and accuracy, were the principal ones of interest to the experiment. However, several other aspects of the experiment were also analysed, and these results will now be reported.

Firstly, performance on the comprehension questions was analysed. As in previous experiments, only accuracy on the two comprehension

questions attached to each experimental passage was scored. The data were analysed as a $2 \times 2 \times 5$ design, with Skill, Foregrounding, and Distance as factors; it was not expected that the latter two factors should have any effect on comprehension accuracy, and this was in fact found to be the case, with both main effects producing F values of less than 1. It was also found that there was no significant difference in comprehension accuracy between the two skill groups; mean accuracy for skilled readers was 2.975, with a maximum possible of 4, and for less-skilled readers was 3.250, with $F(1,22)=2.27$. These results therefore indicate the generally high level of comprehension performance, with no differences between subjects divided on the basis of their reading speeds.

Secondly, the reading times for the intervening foregrounding and backgrounding sentences themselves were analysed. In Experiment 3, it was necessary to analyse these sentences in order to exclude the possibility that some form of spillover effect may have been responsible for the main effect of Foregrounding on pronoun sentence reading time; as Foregrounding had no effect in the present experiment, the main question of interest was whether the Experiment 3 finding, that the foregrounding sentences were read more slowly than the backgrounding sentences, would be replicated. As for Experiment 3, the analysis used only those passages presented at distance 8; in this way, sentence position within the intervening blocks could be considered, with equal numbers of sentences contributing to the mean at each position. The analysis was a $2 \times 2 \times 8$ design, the factors being Skill, Foregrounding, and sentence Position. The analysis was run on the means of the two foregrounded and two backgrounded passages read by each subject. It was found that there was a significant main effect of Skill, with $F(1,22)=8.82$, $p<0.025$; mean reading time for skilled readers was

2191msec, and for less-skilled readers was 2977msec. This shows that reading speed on the full-screen passages successfully predicted sentence by sentence reading speed. There was no main effect of Position, with $F(7,154)=1.13$, and most importantly there was no main effect of Foregrounding, with $F(1,22)=1.73$. The effect found in Experiment 3 was therefore not replicated.

Thirdly, reading times for the first sentences only of the two sets of intervening blocks were analysed; as in Experiment 3, the possibility of interest was the different effect of a topic change in the backgrounded condition compared to topical continuity in the foregrounded condition. For each subject, the mean reading time for the first intervening foregrounding sentence was calculated, taken over all eight passages presented with foregrounding material; and similarly the mean reading time for the first intervening backgrounding sentence was calculated. The means were analysed as a 2 x 2 design, with Skill and Foregrounding as factors. Skilled readers were again found to read faster than less-skilled readers, with $F(1,22)=10.77$, $p<0.005$; however, there was no main effect of Foregrounding, with $F=0$. In this experiment, therefore, subjects do not appear to have responded differently to the two types of intervening material.

Finally, six distractor passages were analysed with regard to the effect of sentence serial position. The previous two experiments found a very long reading time for the first sentence of a passage, with a steady decline to an eventually stable level; in order to generalise this result, six new passages were analysed on this occasion. As the previous two experiments had employed the longest filler passages available, the ones chosen for analysis in this experiment were somewhat shorter. The passages used were all distractors, with three being

selected from each sequence; the shortest had fifteen sentences, and any sentences beyond this in longer passages were ignored. Reading times were analysed as a 2 x 15 design, with Skill and Position as factors; the analysis was run on the means of the fifteen sentence positions over the six distractor passages, for each subject. On this occasion, the main effect of Skill did not reach full significance, with $F(1,22)=3.02$, $p<0.1$; less-skilled readers were approximately 800msec slower than skilled readers. The main effect of Position, however, did prove significant, with $F(14,308)=4.37$, $p<0.001$. The pattern of the effect was identical to that obtained in previous experiments; the first sentence had a mean reading time of nearly seven seconds, while the second sentence had a mean reading time of less than half this figure. Reading times then fell to around the sixth serial position, where they approximately stabilised at between two and a half and three seconds.

Discussion

The main results of this experiment have clearly failed to show any evidence that foregrounding has an effect on recognition memory test performance. In contradiction of the predictions of the working memory theory, backgrounded and foregrounded antecedents were recognised equally rapidly and equally accurately. This outcome must therefore cast doubt on the view that the antecedents shown to be more accessible to pronominal reference in Experiment 3 were being held active in working memory throughout the intervening material.

One problem with criticising the working memory theory on this

basis is that it involves building a conclusion on a failure to obtain an effect, rather than on an effect which has actually been observed. Predicted effects may fail to materialise for a number of reasons, based in faults of experimental design or procedure; such a failure must therefore be accepted cautiously, and measured against the possibility that the experiment may simply have been insensitive. In the present experiment, the important possibility is that the foregrounding manipulation may not have been as effective as in Experiment 3. Although the passages used in the two experiments were substantially the same, a number of slight modifications had nonetheless been made, and in addition the experimental procedure and the nature of the task were considerably different. While there is no specific reason for thinking that these changes may have reduced or eliminated the foregrounding effect, it is a possibility which must be taken into account. This experiment contains no demonstration of the success of the foregrounding manipulation.

In addition to the possibility that the foregrounding effect may have been weakened in this experiment, there may have been other factors contributing to a general insensitivity of the recognition tests. For example, most subjects in their debriefing comments indicated that they considered the comprehension questions to be the crucial task; the recognition tests were rare, and did not require memory for as much information as the questions, while the instructions had not greatly emphasised the recognition responses. In particular, since speed of responding had not been emphasised, sensitivity to effects on latencies may have been impaired. Given the high level of recognition test accuracy, sensitivity of this measure may have been impaired by a ceiling effect. Other causes of insensitivity may have been the lack of practice on the recognition task, and possible motor delays in

responding; with only one test per passage, subjects were unlikely to keep their hands poised over the response buttons. While these are only speculations, it is clear that insensitivity of the recognition tests may be one explanation for the lack of effects.

It should be noted that this experiment also failed to produce any effect of distance. This failure is interesting, since tests at distance 0 should have been facilitated by residual activation of surface traces, as Fletcher (1981) had found. However, the absence of any facilitation at distance 0 may be explained by considering the structure of the antecedent sentences. On most occasions, the crucial antecedent noun occurred very early in the sentence, frequently as the first word and with a clause boundary intervening between it and the end of the sentence. Since, as Chang (1980) showed, it is only the final clause of sentences which remains activated after the sentence has been read, recognition tests at distance 0 may not on average have greatly benefited.

In summary, this experiment has produced results inconsistent with the activation or working memory theory of foregrounding. Such a discovery could be of considerable importance to theories of pronoun interpretation. However, given the dangers of insensitivity noted above, and the fact that the result is based upon absence of an effect rather than on a demonstrated effect, it would be unwise to accept this outcome without replication. Experiment 5 was therefore designed to test this result a second time, with greater safeguards against the dangers of experimental insensitivity.

Introduction

Experiment 5 was carried out with two purposes. While the discussion of Experiment 4 pointed out the danger that the foregrounding effect in that experiment may have somehow been weakened, and that this may have been responsible for the lack of effects on recognition memory, the general validity of the foregrounding effect was not called into question. The discussion assumed that topical continuity as manipulated in Experiment 3 was indeed the source of the pronoun reading time results, and that such results could again be obtained with such a manipulation. However, an effect of such importance should not be uncritically accepted on the basis of one experiment alone. In order to demonstrate the validity of the effect, and its dependence on the particular text manipulation used, it would be desirable to replicate the result both with different subjects and different materials. The present experiment set out to perform the first of these tasks, by using the experimental texts from Experiment 3, as slightly modified for Experiment 4, with a new group of subjects. The appearance of the foregrounding effect on pronoun sentence reading times on a second occasion would strengthen confidence in its genuineness. A full replication of the effect, using new materials as well as new subjects, will be presented in Experiment 6.

The replication of the Experiment 3 foregrounding effect was thus in itself a question of some interest here. In addition however, inclusion of the pronoun sentence test was intended to strengthen any conclusions based on the recognition memory test. If the recognition

test once again failed to show any effect of foregrounding, the presence of a foregrounding effect on pronoun sentences would indicate that this failure was not due to the weakness of the manipulation. Rather, such a failure could be taken as strong evidence against the view that foregrounded antecedents are held active in working memory. This experiment therefore tested for a foregrounding effect on both pronoun sentences and antecedent memory tests, using the same experimental materials and the same subjects for each.

The materials used here in fact constituted a subset of those used in Experiments 3 and 4. Rather than being a complete replication, this experiment only included distances 2 and 8. The comparison between these two distances was considered the most important one, since the main conclusion from Experiment 3 regarding distance was the lack of any additional effect of increasing the number of intervening sentences; use of distances 2 and 8 was sufficient to submit this finding to a second test. In Experiment 3, an interaction between distance and foregrounding was found because the two foregrounding conditions produced equal reading times at distance 0, with only backgrounded reading times being increased when there were intervening sentences; since distance 0 was not included in the present experiment, no interaction between distance and foregrounding was predicted. Details of the texts selected for inclusion in the experiment are given in the Method section.

Finally, it should be noted that this experiment also set out to correct some of the possible causes of recognition test insensitivity in Experiment 4. These changes included giving subjects more practice on the recognition tests, and changing the emphasis of the instructions given; these and other changes are detailed below. Failure to obtain

effects on recognition memory, despite these changes, was expected to be less likely to be due to insensitivity of the test procedure.

Method

Subjects

The subjects used in this experiment were twenty-four first-year psychology students at Plymouth Polytechnic, who received course credit for their participation. They were divided into skilled and less-skilled reader groups after the experiment was completed.

Materials

Sixteen experimental passages were used on this occasion; these were selected at random from the twenty passages used in Experiment 4, with passages 5, 9, 11 and 14 being omitted (these numbers refer to Appendix 6). In order to allow the pronoun sentence reading time technique to be used again, the following changes were made:-

(1) The original pronoun target sentences, first used in Experiment 3, were added back to the appropriate passages. As it was also of interest to confirm the lack of a foregrounding effect on pronoun question accuracy, the original pronoun questions were also reintroduced.

(2) For each experimental passage, a second recognition test word was chosen, in addition to the crucial antecedent noun. These test words were always News, some being plausible and some noncase. The words are

given in Appendix 6.

There were eight conditions in the experiment. Each passage could be varied so as to be presented either for target sentence reading time or for antecedent recognition; with either 2 or 8 intervening sentences; and with the intervening sentences drawn from either the foregrounding or backgrounding block. When a passage was to be presented for target sentence reading time, the Experiment 3 pronoun sentence was inserted after the intervening material, and the new recognition test word was used; when the passage was to be presented for antecedent recognition, the target sentence was omitted and the text was presented exactly as in Experiment 4. The questions used in Experiment 4 were changed only if the passage was presented for target sentence reading time; in that event, the first Experiment 4 question was replaced by the Experiment 3 pronoun question. The second question on each passage was never changed.

In order to reduce the running time of the experiment to one hour, it was decided to use fewer concealed practice and distractor passages than on previous occasions. Only one concealed practice passage was in fact used, being chosen at random from all the filler passages available; this was the fifth distractor passage given in Appendix 3. Two distractor passages for this experiment were also randomly chosen, these being the first and third concealed practice passages given in Appendix 3. Both these chosen distractor passages had New recognition test words; it should be noted, however, that since the experimental passages presented for target sentence reading time also had New test words, there was an ample mixing of tests requiring the two responses, as well as of texts with and without final pronoun sentences.

From this set of materials, twenty-four passage sequences of nineteen passages each were constructed. Each sequence began with the concealed practice passage, which was then followed by a random ordering of the sixteen experimental passages and the two distractor passages. In each sequence, the experimental passages were randomly assigned to the eight conditions, with two passages in each condition. The randomisations were done separately for each sequence, and each distinct sequence was presented to one subject.

One of the intended procedural changes from Experiment 4 was to give subjects more practice at recognition test responding. To this end, twelve new short texts were prepared for use as explicit practice materials. These were two or three sentences long, and consisted of excerpts from other concealed practice and distractor passages used in previous experiments. Some slight modifications were made, and the texts are given in full in Appendix 7. None of these short passages had comprehension questions. Of the twelve recognition test words, half were Old and half New, and these also are given in Appendix 7.

The two reading speed passages used in previous experiments were again employed here to assess subjects' reading ability. These were always interpolated, in the same order, after the fourth and sixteenth texts in the nineteen-passage sequences. The recognition test words used in Experiment 4 were again employed.

Procedure

Each subject attended a single one-hour session, in which a complete passage sequence was read. The passages were presented exactly

as in previous experiments, and the word recognition tests took the same format as in Experiment 4.

The main procedural change in this experiment was in the instructions concerning the recognition tests. These changes were intended to emphasise the importance of the recognition tests, to stress speed as well as accuracy, and to encourage uniformity in the method of responding. Subjects were given the same general instructions as in Experiment 4; however, they were also told to respond as quickly as possible as well as trying to be as accurate as possible, and to move their hand into position near the response buttons as soon as the warning asterisks appeared. The importance of a high speed of responding, consistent with accuracy, was in fact mentioned three times. Subjects read through the twelve explicit practice passages in the presence of the experimenter; during this training, they were again directed to ensure that their hand was resting on or near the response box as soon as the asterisks were displayed. Before being left alone to read through the experimental sequence, subjects were reminded that these passages would be longer than those read in practice, and that they would also have two comprehension questions.

A final procedural change from Experiment 4 was in the recognition latency timings. A new computer technique was devised which allowed latencies as short as 22msec to be recorded; although the limit of about half a second in the previous experiment did not seem to have been a problem, it was felt desirable to use this new method, given that speed was now being emphasised with the expectation of producing faster latencies.

Results

The twenty-four subjects were divided into skilled and less-skilled reader groups in exactly the manner described for the previous three experiments. The product-moment correlation between the two reading speed measures on this occasion was +0.748 (df=22, $p<0.005$).

Main Analyses

This experiment included two types of test on the same materials and the same subjects, involving pronoun sentences and recognition memory responses respectively. Data from the two tests were analysed separately. The pronoun sentence results will be described first, followed by the recognition test results.

Table 5.1

Experiment 5: Mean Pronoun Sentence Reading Times (msec)

	<u>DISTANCE</u>	
	<u>2</u>	<u>8</u>
<u>FOREGROUND</u>	2692	2625
<u>BACKGROUND</u>	3473	3792
	3082	3208

Table 5.1 presents the mean pronoun sentence reading times for the foregrounded and backgrounded conditions, and for distance 2 and distance 8; reading times were analysed as a 2 x 2 x 2 design, with Skill, Foregrounding and Distance as factors. As there were two

observations per condition for each subject, the analysis was conducted on the means of these two values; the ANOVA summary table is given in Appendix 18. On this occasion, it was found that the main effect of Skill on reading time was not significant, with $F < 1$, suggesting that the reading speed measure failed to predict sentence by sentence reading speed as effectively as in previous experiments. All interactions involving the Skill factor also produced F values of less than 1. With regard to the other experimental effects, in Table 5.1 it can be seen that the mean reading time in the foregrounded condition was 2658msec, while the mean in the backgrounded condition was almost a second longer, at 3632msec. This difference was highly significant, with $F(1,22)=21.52$, $p < 0.001$. Reading times in the two distance conditions also differed slightly, but this effect did not approach significance, with $F < 1$. The Foregrounding x Distance interaction was also not significant, with $F(1,22)=1.36$.

The analysis of reading times has therefore confirmed the findings of Experiment 3, showing that pronoun sentences are read faster in the foregrounded condition than in the backgrounded condition and that additional intervening distance has no extra effect beyond distance 2. Answers to the pronoun questions were also analysed, and accuracy on these is given in Table 5.2. The ANOVA summary table is given in Appendix 18. As with the analysis of reading times, the main effect of Skill and all interactions involving that factor produced F values of less than 1. It can be seen from Table 5.2 that pronoun question accuracy was slightly greater in the foregrounded condition than in the backgrounded condition; however, this effect was not significant, with $F(1,22)=1.29$. The main effect of Distance, and the Foregrounding x Distance interaction, were also nonsignificant, with both $F < 1$. These findings again broadly confirm the results of Experiment 3.

Table 5.2

Experiment 5: Mean Pronoun Question Accuracy (max=2)

	<u>DISTANCE</u>		
	<u>2</u>	<u>8</u>	
<u>FOREGROUND</u>	1.625	1.708	1.667
<u>BACKGROUND</u>	1.583	1.458	1.521
	1.604	1.583	

Having confirmed the effects of Experiment 3 on the pronoun sentence data, the recognition test results will now be described. Table 5.3 presents the mean recognition test latencies in the four experimental conditions. As in Experiment 4, the analysis was carried out using only those latencies associated with correct responses; with eight responses per subject, a total of 192 responses was recorded, and only ten of these were errors. The data were analysed using the GENSTAT statistical package which estimated values for these deleted data points. The ANOVA summary table for this analysis is given in Appendix 18, and as for Experiment 4 includes the Trial factor. It was found that for the main effect of Skill, $F(1,22)=1.90$; this was not significant, and similarly all interactions involving Skill proved nonsignificant. In Table 5.3, it can be seen that latencies in the foregrounded condition were approximately 240msec slower than in the backgrounded condition; this difference was in the opposite direction to that predicted, but did not prove significant, with $F(1,22)=1.28$. There was a larger difference in response times between the two distance conditions, with mean latency at distance 2 being 1478msec, and that at distance 8 being 1971msec; this difference, however, failed to reach significance, with $F(1,22)=3.29$, $p<0.1$. The Foregrounding x Distance interaction was not significant, with $F(1,22)=1.47$.

Table 5.3

Experiment 5: Mean Recognition Test Latencies (msec)

	<u>DISTANCE</u>		
	<u>2</u>	<u>8</u>	
<u>FOREGROUND</u>	1451	2240	1846
<u>BACKGROUND</u>	1505	1702	1604
	1478	1971	

The recognition test accuracy data were also analysed, and the ANOVA summary table is given in Appendix 18. As with Experiment 4, there was a very high level of accuracy overall, with only 5.21 per cent errors, and the analysis revealed no significant effects whatever. It was found that for the main effects of Skill, Foregrounding and Distance, and for all interactions, $F < 1$; and in particular, for the Skill x Foregrounding interaction, $F(1,22) = 0.65$. The significant Skill x Foregrounding interaction found in Experiment 4 was therefore not replicated, while the general lack of other effects on recognition performance was fully confirmed.

Subsidiary Analyses

In general, this experiment has therefore confirmed the presence of the effects found in Experiment 3 on pronoun data, and the lack of effects found in Experiment 4 on recognition data. These were the principal questions of interest, but various other analyses which were also conducted will now be reported.

In order to assess general comprehension accuracy in the

experiment, performance on the second comprehension question attached to each experimental passage was scored. The second questions only were scored, as these were presented to all subjects in all conditions; the first question varied according to whether the passage was presented with or without a pronoun sentence. The data were analysed as a $2 \times 2 \times 2$ design, with Skill, Foregrounding and Distance as factors; as expected, there were no effects involving the latter two variables, and in addition the main effect of Skill was not significant, with $F < 1$. Mean accuracy for skilled readers was 1.563 (maximum=2) and for less-skilled readers was 1.646; the level of accuracy was therefore acceptably high, with skill groups showing no differences in performance.

As with the previous two experiments, reading times for the intervening sentences themselves were also analysed. This was especially of interest on this occasion, as the previous two experiments were in disagreement; Experiment 3 found that intervening foregrounding sentences were read slower than intervening backgrounding sentences, while Experiment 4 found no difference between them. As before, only those passages presented at distance 8 were included in the analysis, to ensure equal representation of the eight sentence positions; reading times from both pronoun sentence and recognition test passages were included. In this experiment, therefore, there were four passages in each of the foregrounding conditions at distance 8, rather than two only as in previous experiments. The analysis was conducted on the means of these four passages in each condition for each subject, and constituted a $2 \times 2 \times 8$ design; factors were Skill, Foregrounding and sentence Position. It was found that there were no significant main effects or interactions. The lack of an effect of Foregrounding, with $F(1,22)=1.12$, confirms the finding of Experiment 4, while the lack of a

significant effect of Skill, with $F(1,22)=1.85$, confirms the earlier finding in the present experiment on the analysis of pronoun sentence reading times. A second analysis was conducted on reading times for the first sentences of the two types of intervening block; as in previous experiments, the possibility of interest was that the first backgrounding sentence might be read slower than the first foregrounding sentence, as it constituted a topic change. In this analysis, reading times from the first intervening sentences of passages presented at distance 2 were also included; the 2×2 analysis, with Skill and Foregrounding as factors, therefore used the mean of eight reading times as data points for each subject. It was found that the main effect of Skill was again not significant, with $F(1,22)=1.18$, while for the main effect of Foregrounding, $F<1$. In this experiment, therefore, there were no differences in reading times for the foregrounding and backgrounding sentences.

Previous experiments had found an interesting effect of sentence serial position within a passage; this effect was again investigated. On this occasion there were only two distractor passages in the experiment, the shortest of which had only thirteen sentences; the analysis was therefore conducted using reading times from both these passages, on the first thirteen serial positions only. One of the passages had been used in the serial position analysis of Experiments 2 and 3, while the other had never been used for this purpose before. The analysis was a 2×13 design, with Skill and Position as factors. The main effect of Skill did not reach significance, with $F(1,22)=3.69$, $p<0.1$; mean reading time for less skilled readers was 800msec slower than for skilled readers, but the failure to meet conventional significance levels is further confirmation of the relative weakness of the skill division in this experiment. By contrast, the main effect of

Position was highly significant, with $F(12,264)=7.08$, $p<0.001$. The effect was somewhat more irregular than on previous occasions, as would be expected with fewer sentences at each serial position to equalise random variations in sentence length, but the usual pattern of a very long reading time for the first sentence, and a gradual decline to an eventually stable level at about the sixth serial position, was again evident.

The final analyses to be reported involved comparisons between this experiment and Experiment 4. The instructional changes in this experiment had been expected to produce generally faster recognition latencies than in the previous experiment; in order to test this, a simple comparison was run. For each subject in Experiment 4, a single mean recognition latency was calculated, based on all 44 passages read; this figure therefore included the recognition tests on concealed practice and distractor passages, as well as on the twenty experimental passages. Similarly, for the present experiment a mean latency was calculated for each subject over all nineteen passages read; latencies to the two full-screen passages were not included on either occasion. These two sets of latencies were then compared by means of a t-test; in Experiment 4, mean latency was 2480msec, and in the present experiment mean latency was 2002msec. As this difference was in the anticipated direction, a one-tailed test was used; with 22 degrees of freedom, $t=1.99$, $p<0.05$. Thus, response speeds were successfully improved by the instructional changes between the two experiments. Mean accuracy in the two experiments was also compared, although as no prediction had been made concerning this measure a two-tailed test was used. In Experiment 4, mean accuracy was 0.896, while in the present experiment mean accuracy was 0.942; for this comparison, $t=-3.076$, $p<0.01$. Thus, the increased speed of responding in the present experiment did not take

place at the expense of reduced accuracy levels, with accuracy in fact also being improved.

Discussion

This experiment has produced two principal findings. Firstly, the pronoun sentence reading time results successfully replicated those found in Experiment 3; reading times were faster in the foregrounded condition than in the backgrounded condition, while increasing the number of intervening sentences produced no additional effects on reading time. Thus, the validity of the foregrounding effect is to some extent confirmed. Secondly, the recognition memory test results replicated those of Experiment 4; neither latencies nor accuracy showed any effect of the foregrounding manipulation. Since this outcome was based on the same passages as had shown a foregrounding effect on reading times, the repeated failure of a recognition effect to materialise seems likely to be a genuine reflection of its insensitivity to the foregrounded status of test items.

The confirmation of the foregrounding effect on reading times is only a partial one, as this experiment used passages which had already been used in Experiment 3; however, it is nonetheless important to have demonstrated the effect a second time, using different subjects, a reduced set of experimental materials and an experimental procedure containing new task demands in the form of the added recognition tests. The foregrounding effect of topical continuity thus appears to be robust with regard to particular experimental conditions. In addition, the foregrounding effect in this experiment is free of one possible

alternative explanation which was present in Experiment 3. In that experiment, reading times for the intervening blocks of foregrounding sentences were significantly longer than those for the intervening backgrounding sentences, and it was suggested that this may have made the subsequent target sentences easier to integrate for reasons unconnected with the critical pronoun itself; in the present experiment, the foregrounding effect was obtained despite there being no differences in reading times for the intervening material. Thus, it seems likely that the foregrounding effect in Experiment 3 can also be safely attributed to pronoun interpretation processes.

The foregrounding effect on pronoun sentences gives added confidence to conclusions based on the absence of such an effect on recognition memory. The lack of a recognition memory effect cannot here be attributed to possible weakness in the foregrounding manipulation; and since the robustness of that manipulation has been demonstrated, it seems likely that the foregrounding variable was equally strong in Experiment 4. Furthermore, the suggestion that the recognition tests themselves may have been an insensitive measure is made less plausible owing to the performance improvements found here. The various changes in procedure, amount of practice and emphasis in the instructions were intended to increase performance levels, especially with regard to latencies; the fact that subjects responded significantly faster in this experiment than in Experiment 4 should have increased the sensitivity of the test.

There are, however, two remaining factors which may have led to recognition test insensitivity. Firstly, the instructions given to subjects in the present experiments emphasised the fact that they should decide whether or not the exact test word had occurred in the preceding

passage. From the published studies referred to earlier, it is not possible to tell whether a similar emphasis on exactness was used. The possibility exists that subjects may have delayed their recognition responses while making careful decisions regarding the exactness of their choice; such conscious decision processes may have swamped any effect of activation level. Secondly, in the present experiments a delay of one second was imposed between the end of each passage and presentation of the test word; in the other studies referred to, which did find various significant effects on recognition, the test items were presented immediately after each text. However, it seems unlikely that such a pause would lead to significant decay in item activation, and imposing such a delay in these experiments was intended to ensure rapidity of responses once the test was presented, with no hesitations while making the adjustment from normal reading to recognition decision making. It should be noted that the passages used in these experiments were substantially longer than those used elsewhere (Chang, 1980, for example, used single sentences), while the greater variation in text length was also expected to make recognition tests difficult to respond to if they were not signalled in advance. Nevertheless this is a noticeable difference between these experiments and other published studies; possibly a replication would be desirable, without such a delay being imposed.

These experiments were intended to test the prediction, based on the working memory theory of foregrounding, that recognition memory test performance would be superior if the test item were foregrounded rather than backgrounded. The results of both experiments have clearly shown this not to be the case. Since recognition memory was found to be unaffected by foregrounding, it appears that foregrounded antecedents are not in fact maintained in an activated state as the working memory theory proposes. Although the conclusion that foregrounded items are not more highly activated than backgrounded items is necessarily based on the absence of an effect rather than its presence, this lack of effect has been demonstrated in two separate experiments, and as one of these experiments successfully replicated the foregrounding effect on pronoun sentences the lack of effect on recognition can be accepted as reasonably established.

It is interesting that these direct tests of the prediction have shown there to be no relationship between foregrounding and recognition memory, while the two studies described in the introduction to Experiment 4 (Fletcher, 1981; McKoon and Ratcliff, 1980) had produced results suggesting that a foregrounding effect might have been found. However, there are significant differences between those experiments and the foregrounding manipulation used here. McKoon and Ratcliff (1980) found that concepts were reactivated if concepts linked to them earlier in the text were mentioned again; since the present foregrounding manipulation involved using intervening sentences which either did or did not deal with topics related to the antecedent, this might have been expected to ensure continued activation in a similar way. In the McKoon

and Ratcliff study, however, the relevant linked concept was always part of the same underlying proposition as the reactivated concept, while in the present experiments no such precise analysis was used; and in addition, McKoon and Ratcliff used explicit repetition of the propositionally linked noun in order to produce reactivation of the unmentioned concept, while in the present experiments the foregrounding sentences dealt more generally with features of the scene or actions in which the crucial antecedent was involved. Thus, while McKoon and Ratcliff (1980) demonstrated activation of a concept if a noun linked to it in the same original proposition was explicitly mentioned, this need not imply similar reactivation in cases where only concepts more distantly related to the antecedent, on the basis of a general topical connection, are being explicitly referred to.

With regard to the effect found by Fletcher (1981), it is interesting to note that the strategy used to predict buffer contents was based at least in part on selection of superordinate propositions in the text hierarchy. Since it is known that higher level propositions tend to be reproduced with greater probability in free recall tasks (Kintsch, Kozminsky, Streby, McKoon and Keenan, 1975), it is possible that the effect on recognition memory was caused by similar processes. Thus, higher level propositions may be recognised faster because of the greater strength of their initial coding, or because they are used to organise the text, rather than because they are currently activated at the time of testing. This suggestion implies that the recognition test procedure may be subject to influences on speed of accessing an item in long-term memory, and that a recognition latency difference need not reflect the fact that one item is more highly activated than another. There is some evidence from the present experiments that subjects carry out searches of long-term text memory when making recognition decisions,

and that the influences on these searches are different from those on pronoun assignment. In Experiment 4, it was noticeable that while recognition latencies at distance 0 to 6 were fairly constant, latencies increased greatly at distance 8; this effect did not approach significance, but in Experiment 5 the latency increase from distance 2 to distance 8 did approach significance. There is thus some indication that test items are harder to recognise the further back in the text they had occurred. This would be a plausible result, and since pronoun sentence reading times are clearly unaffected by distance it supports the view that the processes involved in pronoun assignment are not similar to those used in recognition test decisions.

The failure to obtain any effect of foregrounding on recognition memory is clearly incompatible with the working memory theory of pronoun interpretation. On that theory, foregrounded antecedents are seen as being already activated in working memory, and should thus be more accessible to recognition tests; no evidence in support of this was found. Nevertheless, it is clear that foregrounded antecedents are readily accessible to pronominal reference. In the light of the present results, it is evidently necessary to consider whether some explanation of the foregrounding effect can be framed which does not rely on working memory presence of antecedents. The following chapter addresses this question.

Before continuing, however, an important qualification must be made regarding the conclusions drawn from Experiments 4 and 5. These conclusions evidently rest on a failure to obtain an effect rather than on an actually observed effect; and it was felt necessary to replicate this failure in order to strengthen confidence that no foregrounding effect on recognition memory in fact existed. However, there are various reasons why the repeated failure to obtain such an effect need not constitute strong evidence against the predictions of the working memory theory. Firstly, the recognition test data obtained here consisted of only two observations per condition per subject; it is possible that this small number of observations may have allowed genuine effects to be obscured by noise in the data. The fact that the large latency difference between distances 0 and 8 found in Experiment 4 did not approach significance would be consistent with this view, and it is similarly possible that any foregrounding effect on recognition memory may have been obscured by noise. Secondly, as noted above, the recognition procedure used in these experiments was somewhat different from that used in various published experiments which have succeeded in obtaining effects; it is possible that these differences reduced test sensitivity to such an extent that no effects whatever could have been detected. It is striking that the only significant recognition effects found here were the inter-experiment comparisons attributed to instructional changes. Thirdly, the large foregrounding effect obtained on pronoun sentence reading times need not imply the presence of a similarly large effect on recognition memory; although the use of only two observations per condition proved sensitive to the reading time effect, it need not be sensitive to any effect on recognition tests.

For these reasons, it can be argued that the lack of a foregrounding effect on recognition memory has not been securely established by these experiments. The recognition memory predictions based on the working memory theory of foregrounding require a more thorough investigation, obtaining more observations per condition per subject in order to reduce possible effects of noise, and using a test procedure more closely comparable to that described in the published papers referred to above.

CHAPTER IV

Alternative Explanations of Foregrounding

- 1 Criticisms of the Working Memory Theory
- 2 Retrieval Processes and Foregrounding
- 3 Conclusion

The fundamental objective in studies of pronoun interpretation is to understand the processes by which a text is integrated into a unified whole. Forming connections between pronouns and their antecedents is one part of this integration, and investigations of foregrounding attempt both to identify the influences on pronoun assignment and to develop theories to explain how they operate. The type of theory which has been widely accepted proposes that a certain rapidly accessible partition of memory, which is being referred to here as working memory, is used to hold representations of entities which for textual reasons seem to be important and hence likely to be mentioned again. If the writer uses pronouns appropriately, the reader will be able to connect them directly to this store. Experiments 1-3 were carried out within the framework of such a theory, and the results obtained there appear to be compatible with an explanation of this type. Experiments 4-5, however, made a direct test of one of the predictions of the working memory theory, and failed to produce any evidence in support of it. In this section, therefore, the motivation for and assumptions behind the working memory theory will be more closely examined. It will be argued that the working memory theory imposes unnecessary restrictions on pronoun interpretation processes, and that the available evidence is equally compatible with a different class of theory which would be able to predict the lack of foregrounding effect on recognition memory.

The principal motivation for the working memory theory lies in the apparently formidable nature of the assignment problem. When a reader encounters a pronoun, full comprehension and integration cannot be attained until that pronoun is assigned to a referent; in the

circumstances under investigation here, the correct referent will be an entity mentioned earlier in the text. Thus, the assignment process requires some form of memory search in order to determine the correct antecedent. However, even in relatively short texts the number of entities already mentioned, and the complexity of the memory representation in which they are embedded, may be very great; and it is argued on these grounds that assignment processes requiring a search of the entire representation could not be carried out within a reasonable time consistent with normal reading speeds. In order to facilitate assignment, and avoid prolonged delays in reading, the search area for pronoun antecedents must be limited in some way.

This argument can be criticised on a number of grounds. In general, it clearly depends on an assumed limitation to human comprehension processes, rather than on a demonstrated limitation. No justification is normally given for claiming that human cognitive abilities would be inherently unable to cope with the assignment problem as formulated above; rather, the assertion appears to have been based on the needs of computer programs in artificial intelligence research, which do require a limited search area in order to resolve pronominal references efficiently (e.g. Grosz, 1981; Sidner, 1979). Whether human comprehension processes are similarly lacking in power is open to question. A more specific criticism of the argument, however, concerns the assumed nature of the antecedent search process, and this will now be discussed in detail. For convenience, the basis of this discussion will be the arguments put forward by Sanford and Garrod (1981); this is the most explicit model available in the literature, and contains the most complete exposition of the assumptions at issue, while the overall nature of their theory is compatible with most other treatments of anaphora resolution.

In Sanford and Garrod's model, the antecedent search process in any particular case is specified by the values given to three constraining variables. These variables define the search process in terms of: (a) the partition of memory which is to be searched; (b) the available partial description of what is being sought; and (c) the type of information which a successful search must produce. For an anaphoric pronoun, the partition of memory which is to be searched is restricted to a domain termed explicit focus; this corresponds to working memory in present terminology. The third variable would be specified as requiring some unique identity, or token for an entity; that is, the search process must produce some particular referent. It is the second variable, however, which is crucial. Clearly, in order for any memory search to be carried out there must be some preliminary definition of the target of that search; proposals concerning the nature of the search process depend on what that preliminary definition is assumed to be. For Sanford and Garrod, the partial description used to define the search process for a pronominal referent consists of "the semantic specification of the pronoun itself," which in the case of the pronoun 'he' would amount to no more than "entity, male, singular" (Sanford and Garrod, 1981; page 157). The search process attempts to locate any particular entity which matches this partial description. As Sanford and Garrod point out, with only this definition of the target of search "a search of all memory would be fruitless, since it would return every male entity" (ibid.). Not only would a search of this type require an unreasonable length of time to carry out, it would in fact be inherently unable to produce a satisfactory result. It is therefore the assumption that only very limited information is used as a semantic constraint on the search process which requires that the search be restricted to a narrow domain.

This assumption appears to be a special case of a more general processing assumption, which is brought out in full by Just and Carpenter (1980). To defend their use of local gaze duration as the primary data on which they construct and test their model of reading, Just and Carpenter make the assumption that no aspects of comprehension processing are deferred; they argue that an interpretation of a word is attempted as soon as it is encountered, including attempts to assign it to a referent. Thus, in the case of a pronoun, an attempt would be made to locate an antecedent as soon as the pronoun was read. Just and Carpenter's immediacy assumption is also implicit in the widely used terms such as 'immediate matching' or 'direct mapping'. Sanford and Garrod's more detailed analysis is thus merely a refinement of a generally held assumption, namely that an attempt is made to assign a pronoun as soon as it is encountered and that the only relevant constraints are those contained in that lexical item itself.

It is clear, however, that pronoun assignment is constrained by many more influences than the semantic specification of the pronoun itself. Many of these additional influences were reviewed earlier, and will only be mentioned here. Concerning within-sentence anaphora in particular, it has been shown that assignment may be constrained by aspects of syntactic structure, by stress patterns, and by verb semantics; while there are clearly effects of the nature of the events or actions that are being predicated of the entities involved, and of the relationships between them. Some of these factors could plausibly be explained as influences on which entities were in working memory when the pronoun was encountered; however, the fact that assignment is affected by relationships between pronoun and antecedent such as parallel function (Sheldon, 1974), or by sentence content occurring after the pronoun (Ehrlich, 1980), cannot be accounted for in this way.

Furthermore, it should be pointed out that in the case of pronouns which occur relatively late in a sentence the reader already has potentially relevant information available by the time the pronoun is encountered; it would seem reasonable to allow the antecedent search process to be constrained by the earlier content of the sentence in such cases. The immediacy assumption of Just and Carpenter (1980) would be compatible with this extension to the direct assignment process, but Sanford and Garrod (1981) explicitly assert that mapping takes place in advance of semantic interpretation.

In general, the point being made here is that pronoun assignment involves the active selection of an antecedent, rather than relatively passive acceptance of whatever entity is currently in working memory. Crucially, preferred assignments are frequently dependent on relationships between pronoun and antecedent; they are not dictated in advance by antecedent status alone. The antecedent search process involves complex decisions concerning what is or is not a potential antecedent in any one case, rather than simple direct matching to a predetermined memory location. If it is assumed that the search process is guided or constrained by a partial description of the desired entity which has been elaborated on the basis of other aspects of the sentence than the pronoun itself, then the apparent difficulty in searching a large memory store is considerably reduced. If in particular the target of search has been given an extended definition on the basis of any available semantically disambiguating information, then the number of acceptable entities in the memory representation of the preceding text is likely to be very small. The example passage from Experiment 3 will serve to illustrate this. The target sentence in that passage was 'He stood back at last and shouted for a crowbar', and it is clear that this sentence conveys a great deal of information about the person involved.

In addition to specifying a male individual, the sentence implies that this person is involved in some task which could require the use of a crowbar. The fact that the person 'stood back at last' implies that he has been carrying out the task for some time already, while to shout for something may imply urgency. None of this information uniquely specifies that the person is a fireman attempting to break into a burning building; however, the use of such an elaborated definition would allow the antecedent search process to rapidly decide that such an individual was the desired entity. Thus, even without considering the possible use of syntactic or structural information, a more semantically constrained search process would clearly be able to select the required antecedent from memory. It should be noted, however, that while this argument answers Sanford and Garrod's claim that a search of all memory would "return every male entity", it does not yet explain how such a search can be rapidly accomplished. This issue will be dealt with below.

The suggestion being made here, that the antecedent search process may be guided by much more information than the specification of the pronoun alone, is not in itself controversial. Rather, the point of disagreement between the present suggestion and a standard working memory theory concerns the place which this elaborated information has in the assignment process. Sanford and Garrod (1981) recognise that antecedent searches can occur in this way, but they relegate such processing to a secondary stage which is attempted only if the primary stage of immediate matching fails. Similarly, Just and Carpenter (1980) accept that some referents may not be assigned until the ends of sentences, when all ambiguities have been clarified. It is clear that this view makes a distinction between influences on assignment and the processes by which such assignments are made. Eventual assignment of a

pronoun is accepted as being susceptible to complex semantic and structural influences, but it is argued that these are only taken into account if simpler interpretation processes prove insufficient. The working memory theory thus includes two distinct search mechanisms, with very different characteristics; and these are assumed to operate on two distinct partitions of memory. The above discussion has shown that it is not necessary to postulate such complexity in order to account for successful assignment; the question to be discussed now concerns whether such a distinction between primary and secondary processing is necessary to account for the available empirical evidence. This discussion will revolve around three points; the interpretation of the foregrounding effect on reading times, the need to account for very rapid changes in item availability, and the fact that semantically ambiguous pronouns may nonetheless receive generally agreed assignments.

The principal empirical evidence offered in support of a distinction between immediate matching in primary processing and more elaborated searching in secondary processing is the pronoun sentence reading time difference between foregrounded and backgrounded conditions. This difference is explained by the additional processing which the working memory theory predicts in the backgrounded condition. In both conditions, when a pronoun is encountered the primary processing stage attempts an immediate match to working memory; this proves successful in the foregrounded condition, but fails in the backgrounded condition. Longer reading times in the backgrounded condition are thus explained by the need for further processing after a processing stage common to both conditions has been completed. It may incidentally be noted that this account assumes that failure to complete a direct mapping takes at least as much processing time as successfully

completing a mapping; but it is in any case obviously unnecessary to postulate additive processing stages in order to account for a reading time difference. An equally plausible account could argue that pronoun assignment was achieved by a single processing mechanism in both foregrounded and backgrounded conditions, but that this mechanism was slower to succeed in the latter. This argument will be pursued in detail below; it is enough to point out here that there is nothing in the available reading time data which compels the working memory explanation. Reading time studies show that pronoun sentences are easier to interpret in some contexts than in others, but the source of this difficulty is not revealed by the data in isolation. Furthermore, although the working memory theory predicts a reading time difference between foregrounded and backgrounded conditions, it is interesting to note that attempts to derive and test other predictions based on the theory have not been made. The present Experiments 4-5 carried out such a test, and found results which appeared incompatible with the working memory theory; similar attempts to directly test the processes underlying that theory have not, however, been presented in the literature. Although there are many studies demonstrating how different text features may influence the ease of pronoun interpretation, these demonstrations are in effect simply more instances of the phenomenon which the working memory theory sets out to explain. The foregrounding effect consists of the fact that topical or important entities are easier to refer to than less important entities; studies of this effect have served mainly to establish the influences which determine topicality, rather than to reveal the processes by which topicality operates.

Another aspect of foregrounding which has been said to require some form of working memory explanation is the rapid variation in the

foregrounded status of items which can occur in a text. Sanford and Garrod (1981) point out that different items are constantly gaining and losing foregrounded status as a text is read; clearly, any suggested explanation of foregrounding must be able to account for these rapid changes. It is partly in order to encompass this flexibility that Sanford and Garrod suggest a division of memory into two partitions, one of which is stable and enduring while the other, called focus, is dynamic and rapidly altering in content. The first of these partitions includes the long-term memory representation of the text, while the second, corresponding to working memory, holds the currently foregrounded information. Thus, variations in foregrounding are explained by rapid processes which enter items into focus and clear them from it when required. However, there is in fact no need to postulate a separate dynamic partition of memory in order to account for changes in item availability. This explanation is required only if the antecedent search process is regarded as a relatively simple process of attempted direct mapping based on very limited information; if the search process is allowed to have greater complexity, and to make use of more elaborated target definitions, then rapid changes in availability are made more explicable. The long-term representation of the text is itself far from unchanging. As a text is read, more information is constantly being added, while the structure and organisation of the representation is presumably also being updated. Interactions between a constantly evolving text representation, and a complex antecedent search process, offer many possibilities for explaining both the foregrounding effect itself and the rapid changes in foregrounded status.

Sanford and Garrod (1981) put forward another argument in support of their view that direct mapping between pronoun and antecedent takes place in advance of semantic interpretation. They point out that in

many cases a logically ambiguous pronoun is nonetheless given a generally agreed assignment by all readers; in such cases, an assignment process based solely on semantic disambiguation would be clearly inadequate. However, the fact that a particular assignment is not determinate on semantic or pragmatic grounds does not preclude the use of complex retrieval processes in selection of that assignment; such retrieval processes may select a preferred antecedent, even if unable to guarantee its correctness. In addition, as noted above the antecedent search and decision processes can be viewed as making use of syntactic and structural constraints, as well as semantic ones; in the context of the present series of experiments semantic factors will be emphasised, but the role of other cues must also be allowed for. It is not necessary to assume that a certain antecedent has privileged status which makes it accessible to direct mapping in order to account for agreed assignments of semantically ambiguous pronouns.

The overall trend of the discussion so far can be summarised as follows. Working memory theories of pronoun assignment suggest that there is an initial processing stage involving a severely degraded search process, which is constrained by limited information and is insensitive to features of text structure. In order to allow this search to be successful, it is assumed to be limited to a narrow search domain containing few entities and in which desired antecedents can therefore be easily found. However, if this direct mapping stage should fail, the theory argues that more complex retrieval processes operating on a different search domain are then called into play. This theory was criticised as not being required by the evidence, which could equally be explained by a type of theory which allows complex and more elaborated search processes as the normal method of assignment; details of this alternative explanation are given shortly. The working memory theory

emphasises sensitivity to signals in a text which cue the appropriate selection of working memory contents, but reduces the assignment process itself to a rigidly predetermined sequence of relatively automatic stages; by contrast, the alternative seeks to explain foregrounding phenomena by stressing the elaborate, flexible and adaptive nature of those assignment processes.

The working memory theory is evidently costly in terms of the number of search processes and the number of memory partitions which it claims are involved in pronoun assignment. All versions of the theory accept that complex and extended retrieval processes are eventually resorted to in backgrounded conditions, but the nature and operation of these processes have not been investigated. Nevertheless, the theory proposes that a completely separate and substantially different search process underlies successful assignment in foregrounded conditions. This suggestion must also allow for some form of mechanism which eventually uses semantic and structural information to detect erroneous immediate mappings. In addition, the theory is forced to propose that searches are made of a number of functionally distinct memory partitions; most versions of the theory contain only two such stores, but Sanford and Garrod (1981) in fact subdivide them further and suggest that there are four. It is being argued here that this level of complexity is not required in order to explain foregrounding phenomena, but that a theory proposing only one elaborated search and decision process would be adequate. Such processing is accepted as necessary in the backgrounded case, and it would be parsimonious if these mechanisms could be applied to the foregrounded case as well. It should also be pointed out that similar complex decision making may be required in the foregrounded case even within a working memory theory. This possibility arises because working memory is assumed to contain representations of

several entities, not simply of one foregrounded entity. Direct mappings to working memory would therefore themselves require some form of disambiguation. Sanford and Garrod (1981) resolve this difficulty by allocating different amounts of storage space to different entities, and asserting that the one with the most storage space will be most readily accessible. This added complication not only implies that there may be degrees of foregrounding, when in fact the empirical evidence, such as that contained in the present Experiment 3, suggests that foregrounding is an all-or-none distinction, but also subtly undermines the basis of the working memory theory as a whole. Working memory presence was originally proposed in order to explain direct accessibility; this new suggestion implies that items may be present in working memory and yet remain relatively inaccessible. This contortion is a further complexity of the theory which would make a simpler alternative still more attractive.

The above discussion of the assumptions underlying the working memory theory of pronoun interpretation was prompted by the results of Experiments 4-5, which were apparently incompatible with the suggestion that foregrounded antecedents are held in a directly accessible store. The discussion has shown that the working memory theory is not the only potential explanation for foregrounding phenomena, and that the theory itself contains several unnecessary complexities. An alternative class of theory is discussed in the next section.

Before presenting a detailed description of an alternative to the working memory theory, however, one possible point of confusion must be clarified. The arguments being put forward here may superficially appear to be an attack on the general concept of the use of a working memory system in reading, and to suggest that no such store is required;

however, the aim of these arguments is in fact much more limited. The need for a common processing system within which information can be compared and combined, and where the control functions which direct the reading process are executed, will underlie the present research as closely as it does the working memory theory itself. The question at issue here is not whether a working memory system is used in reading, nor even whether its contents and processes play some role in foregrounding; it will in fact be argued that residual information from the preceding sentence which is still available during processing of the current sentence makes a significant contribution to ease of antecedent retrieval. It is also the case that working memory plays a role in the interpretation of single clauses or sentences, by maintaining earlier words for connection with later input. The question to be considered here concerns solely whether text manipulations which are known to affect ease of pronoun interpretation do so by determining whether or not antecedent items are directly maintained in working memory. The working memory theory of foregrounding argues that foregrounded entities are held in an activated state for possibly lengthy periods, despite not being explicitly mentioned; the retrieval theory outlined below specifically contradicts this suggested function of working memory, but preserves the generally agreed importance of such a system as a whole.

The preceding section presented a critical analysis of the processing assumptions underlying working memory theories of foregrounding, and argued both that present evidence does not compel acceptance of the theory and that such a theory would in fact be cumbersome and excessively complex. The present section describes the general features of an alternative class of theory, and shows how this theory can be made to explain existing knowledge of foregrounding phenomena. This class of theory concentrates on processes of antecedent retrieval, and strategies of pronoun assignment, rather than on prior selection of items for inclusion in working memory. In order to illustrate the ways in which antecedent retrieval can be used to explain foregrounding phenomena, this discussion will be principally concerned with the foregrounding manipulation of topical continuity used in the present Experiments 3-5. This emphasis is to some extent justified by the fact that the remainder of this thesis continues to make use of that manipulation, and attempts to elucidate the ways in which it produces its effects. Furthermore, it should be pointed out that the main purpose of this discussion is simply to demonstrate the type of retrieval processes that are being put forward as an explanation for foregrounding, and to show how they may operate in a particular case; rigorous specification of those processes, with their application to other text manipulations, will not be attempted. Despite this, however, other foregrounding phenomena described in the literature will continue to be referred to at appropriate places, and their connections with the explanation of the topical continuity effect will be noted.

The alternative class of theory of foregrounding, which for brevity will be referred to as the retrieval theory, in fact contains two criticisms of the working memory model. The working memory theory contains two central assertions. Firstly, it is argued that items from a text may be held in working memory for long periods, despite not being explicitly mentioned again, provided that textual cues indicate that they are still important or relevant concepts; and secondly, it is claimed that pronoun assignment consists of an initial attempt to make a direct match to working memory based on a limited specification of the required antecedent. These two components of the theory are logically distinct; as was noted earlier, the presence of two or more entities in working memory may still require that pronoun assignment involve the use of an augmented definition of the object of search. Although the retrieval theory therefore argues that foregrounding phenomena can be explained by extended retrieval processes operating on a structured long-term text representation, it is nonetheless possible that such processes may at some stage also operate on a more limited working memory representation. This possibility will be returned to again in connection with Experiments 6 and 7; for the present, however, the retrieval theory will be described in terms of both augmented searching and absence of antecedents from working memory. The later discussions of this topic will argue that although the two central assertions of the working memory theory are logically distinct, they are in fact both essential if that theory is to be meaningful.

The retrieval theory therefore argues that differences in the ease of pronoun interpretation can be explained by variations in the difficulty of locating the desired antecedent in long-term text memory. As a corollary to this, it is also argued that some of the effects on pronoun interpretation may be due to variations in the difficulty of

deciding or verifying that a located concept is in fact the correct referent. When a pronoun is encountered, the reader executes a complex search and decision procedure involving the use of an augmented definition of the object or target of search, and constrained in operation by features of syntactic and textual structure. Foregrounding phenomena can be explained by interactions between the type of information available to guide the search process and the context of occurrence of pronoun and antecedent; foregrounding is thus seen as a relationship between a concept and a particular occasion of reference to that concept, rather than as a quality determinable by considering antecedent status alone. There are thus several factors which may contribute to overall ease of pronoun interpretation. There may or may not be adequate information to define the nature of the required antecedent; this information may be a useful or less useful retrieval cue to the appropriate subsection of the text representation; or the structural relationship between pronoun and antecedent locations may be complex or straightforward. These suggestions will now be applied to the topical continuity manipulation first used here in Experiment 3.

Firstly, it is evident that differences between conditions in that experiment cannot be attributed to differences in the disambiguating information available from the target sentences themselves, as the same sentences had naturally been used in both the foregrounded and backgrounded conditions. However, it was argued earlier that one of the functions genuinely attributable to working memory may be to carry over information from the previous sentence to assist in integrating the next; this may occur as an explicit integrative process, or it may simply be an unavoidable outcome of the residual retention in working memory of the results of processing the earlier sentence. In whichever way it occurs, it is clear that the information available from the

preceding sentence will be quite different in the foregrounded and backgrounded conditions. In the foregrounded case, the intervening sentences are closely related to the disambiguating content of the target sentence; this necessarily follows, since the crucial pronouns were disambiguated by reference to the aspects of the antecedent which also formed the broad topic area of the foregrounding sentences. In the backgrounded case, however, the intervening sentences bear no relation to the target sentence, since they were explicitly written to deal with topics that had no direct or unique connection with the crucial antecedent. The consequences of this difference are as follows. In the foregrounded case, whatever information is available from the preceding sentence will be consistent with and can be combined with the target sentence content; as such, it is likely to be of value in adding to what is known about the required antecedent. This will thus lead to a more detailed specification of the type of entity it must be or of the type of actions and events it must be involved in. By contrast, in the backgrounded case the residual information from the preceding sentence will be of too general a nature, dealing with features of the story setting or events that are not specific to one individual, to be of any value in specifying the pronoun's referent. Thus, in the foregrounded case the search process can operate with a more fully specified augmented definition.

The foregrounding manipulation of topical continuity thus produces a more detailed and consistent specification of the required antecedent in the foregrounded case. This specification is assumed to be used both as a retrieval cue which directs the search process through the text representation, and as a decision criterion which determines when the correct antecedent has been found. Both these processes will be assisted by having more information available. However, they will also

be assisted by the greater usefulness of that information in the foregrounded case. When the initial reference to the crucial antecedent entity is followed in the backgrounded condition by sentences dealing with other topics, it is evident that the strength and nature of the antecedent's coding in memory will be a function solely of the antecedent sentence itself; the antecedent, and the antecedent's defining attribute, are likely to be treated no differently to the other individuals mentioned in the introductory section. In the foregrounded case, however, the subsequent sentences which deal with topics related to the antecedent will lead to that feature becoming more prominent in the text representation. When, therefore, an antecedent search process is initiated which makes use of that defining feature, a relevant section of memory will prove easier to locate than in the backgrounded case, where no particular subsection of memory has been coded in relation to that attribute. In addition, since more information relevant to that attribute will already have been integrated with the antecedent in the foregrounded case, the decision that this is the correct referent will be made more readily.

These latter suggestions thus argue that in the foregrounded case sections of the text representation related to the antecedent specification will be more easily located, as that specification will be a more important aspect of text memory. This suggestion can be formalised in terms of the construction of semantic macro-structures. As was described in the literature review, the operation of macro-rules serves to reduce the availability of certain pieces of information, but to maintain and increase the availability of others by expressing them in higher-level macro-propositions. Particular facts or details explicitly mentioned in the text may be omitted from the macro-structure so long as they are not required in the interpretation of subsequent

sentences. In the passages used in Experiments 3-5, a plausible location for a macro-structure boundary is at the end of the constant block of introductory sentences; these share the general purpose of establishing the setting of the passage, and describe the various activities of different individuals involved in the scene. In the backgrounded condition, however, the introductory block is followed by sentences which move away from the specific actions of the individuals which have been introduced there, and which deal instead with more generally relevant topics. Thus, the content of the early sentences may be subjected to the various macro-rules, and macro-propositions will be constructed which omit specific detail but give sufficient general information about the setting to allow comprehension of the following sentences; this outcome is the natural result of the operation of the rules termed Deletion and Generalisation by van Dijk (1977a). In the foregrounded condition, however, the macro-rules will operate under a different set of constraints. In that condition, the intervening sentences continue to deal with a topic that is closely related to one of the specific entities mentioned in the introductory block; the existence of that entity, and the nature of the particular actions or events in which it was involved, may therefore be said to constitute interpretation conditions for full understanding of what follows. This constraint will prevent the macro-processes from deleting that information, or from subsuming it under a generalised macro-proposition representing the introductory detail. Thus, while the particular facts relating to the other entities in the passage may be omitted from the macro-structure, the facts relating to the crucial antecedent must be retained. These details may, for example, be carried directly into the macro-structure unchanged, under the macro-rule termed Integration (van Dijk, 1977a).

The consequences of these differences in the operation of macro-rules in the foregrounded and backgrounded conditions have clear relevance to antecedent retrieval processes. The semantic macro-structure is regarded as the basis on which a memory representation of a text is organised, and in addition is seen as acting as a retrieval cue when attempting to recall the text (van Dijk, 1977a). Information contained in the macro-structure is directly available in the representation, but lower levels of information are located both by searching downward through the structure and by inverse application of the macro-rules. Although these suggestions were originally put forward as explanations for features of explicit text recall after complete reading of a passage, it is also true that macro-structures are under continuous construction during reading and should therefore also influence the retrieval processes involved in pronoun assignment. In the foregrounded condition, a retrieval search based on information combined from the target sentence and the preceding sentence will readily locate a macro-proposition bearing a direct semantic relationship to that specification; if the antecedent itself is represented in the macro-proposition assignment will be correspondingly swift, but even if it is not it will be found by downward searching in that specified area of the text representation. In the backgrounded condition, however, not only will the antecedent not be directly available in the macro-structure, but also there will be no macro-proposition of clear relevance to the augmented antecedent specification. This will be so, because macro-processes acting to reduce the content of the introductory block will have deleted both the antecedent and its defining attribute. Thus, the search process will have no match to the retrieval cue which would indicate the optimum area of memory to search. Even if searching were to operate downward from the generalised macro-proposition representing the introductory block,

the specific micro-propositions concerned with the desired antecedent will have been coded or structured very similarly to those representing the other entities, and so will be difficult to locate. By the same token, confirming an unambiguous assignment in the backgrounded condition will also be more difficult, and hence decision processes will be slowed as well as retrieval processes.

These suggestions regarding the effects of macro-structure formation are clearly informal ones, and no attempt will be made here to carry out rigorous propositional and macro-structural analyses of the experimental passages. However, the suggestions are fully consistent with the theory of semantic macro-structures, which as the literature review showed has received substantial empirical support. In addition, the relevance of these general observations would not be reduced if the specific features of the Kintsch and van Dijk model were rejected. The formal theory of macro-structures is an attempt to account for undeniable aspects of text comprehension and memory, such as the ability to identify topics and write summaries, or the selective loss of different types of information. However these empirical facts are represented in psychological theory, processes involving the abstraction of relevant information and the discarding of unimportant details will continue to have a bearing on the pronoun assignment processes suggested here. For example, even if comprehension were seen as the construction of a mental model, it is clear that such models must be selective in content and structured in organisation; whatever form such models may take, it would be implausible to suggest that a complete model representing a long text could be retained in full detail.

The above discussion of antecedent retrieval processes has centred around the consequences of having a more detailed specification of the

object of search in the foregrounded case, and of the availability or salience of a matching semantic structure in the text representation. However, in addition to these effects of semantic constraints on the retrieval process, there may also be important structural constraints which either hamper antecedent location or which conflict with an apparently semantically acceptable assignment. It has so far been assumed that only semantic agreement between pronoun and antecedent was necessary for successful assignment, and that macro-propositions were useful simply as markers of semantically relevant sections of the text representation. There is however also evidence that texts may be given a structural analysis which is independent of its particular semantic content; the principal research in this area has been concerned with analyses based on story grammars, and was briefly reviewed earlier. For example, Thorndyke (1977) presented subjects with passages which had been analysed into tree structures of story categories, such as setting, episode, or resolution; in one experiment, subjects were presented with two passages which had identical plot structure but different semantic content. The important possibilities for present purposes are two-fold. Firstly, retrieval searches for pronoun antecedents may be constrained by the abstract structure of the text; the degree or type of structural separation between pronoun and antecedent may have consequences for ease of retrieval. Secondly, the acceptability of a referent as antecedent for a particular pronoun may be influenced by the structural relationship between them; a concept which semantically matches what is known of the pronoun may nonetheless be rejected if it occurs in an inappropriate section of the text representation. These suggestions will first be illustrated with reference to an abstract analysis of story structure, and will then be applied to the texts used here in Experiments 3-5.

As was described earlier, story structures may be represented in standard tree diagrams, with nodes representing story categories and terminal symbols consisting of the actual text propositions. Such a schema has effects on recall, but is also relevant to comprehension as readers use various types of cue in the text to identify categories and transitions between categories. Processing is sensitive in different ways to story structure. For example, readers may know what type of information to expect while processing a category representing a story complication, or may anticipate connections between a setting and a resolution. Possible effects on pronoun assignment are as follows. Firstly, in well-constructed texts it may be expected that pronoun and antecedent will most probably occur within the same structural category; this need not be the category represented by the most immediately superordinate node, but may be a higher-level node dominating both the current category and other related categories. This suggestion receives some support from the observation by Mandler and Goodman (1982), that in writing their experimental texts they tended not to include pronouns in the first sentences of story constituents. Thus, the search process may be directed first to concepts represented within the current structural block, or to those branching downward from a shared higher category. Secondly, the duration of the search process may be a function of the number of structural nodes separating the occurrence of the pronoun from its antecedent; although the available specification of the antecedent may indicate a particular subsection of memory, not all such areas will be directly accessible. In particular, difficulties may be experienced when accessing information contained in a branching structure which does not share a common dominating category with the current branch of the tree representation. Such difficulties may occur not only because of the added time required to traverse the intervening nodes while executing the search, but also because the search process is initially

prevented from crossing major structural boundaries; failure to locate an antecedent within the most relevant structural category may lead first to a more exhaustive search of that category, before the search is finally directed to structurally more improbable locations. Thirdly, and related to this, a semantic match between pronoun and antecedent of a given exactness may be more readily accepted if the two occur within a single structural category; because of the structural reasons for uncertainty in an assignment across major category boundaries, a stronger criterion for semantic acceptability may be applied. Thus, structural incompatibility may be countered by semantic closeness. In this way, decision processes as well as search processes may be made sensitive to structural factors.

These suggestions concerning the influence of text structure are clearly as informal as those regarding the effects of macro-structure content, and are also more speculative since they have not been related to any specific model of story grammar. In addition, the theory of story structure has not been especially applied to questions of referential coherence, while the Kintsch and van Dijk model of macro-structure formation places a strong emphasis on such interconnections between propositions. Nevertheless, these structural considerations can be applied to the texts which have been used here, although rigorous analysis will again not be attempted. The most striking feature of those texts is that in the foregrounded case topical continuity between the antecedent sentence and the eventual pronoun sentence must lead to them both being dominated by a single category node; this node may not dominate both sentences directly, but it is evident that virtually any formal analysis must place them in a single structural block. By contrast, in the backgrounded condition the intervening sentences are unconnected with the specific topics of the

introductory block, and may readily be construed as the development of a new and structurally distinct subsection of the passage representation. Thus, when the target sentence is encountered in the backgrounded condition, the pronoun and antecedent sentences are likely to share no common category other than that of the passage itself. In this way, antecedent retrieval processes will be assisted by the structural constraints on their operation in the foregrounded case, but hampered by them in the backgrounded case.

One reason for not attempting a more detailed analysis of the experimental texts used here is that they do not readily fall into any text type which has received such an analysis in the literature. Those analyses have generally been applied to stories taken from a highly conventionalised story type, such as folk tales or simple adventure stories; the passages used here are not comparable to such schema types. However, the suggested relevance of story structure to anaphora resolution is clearly open to empirical investigation, and the use of more stylised text types would allow precise study of the effects of such variables as the number and type of category nodes intervening between pronoun and antecedent. The present observations are sufficient to indicate the wide range of factors which must be taken into account when considering the operation of possible antecedent retrieval processes, and to show that such processes would be sufficiently malleable to account for differences in the ease of pronoun interpretation.

The foregrounding effects found in the present Experiments 3 and 5 are not the only ones which can be related to this account of the influences of text structure on antecedent retrieval. Various other findings in the literature may also be explained in this way. Firstly,

it is evident that the manipulation used in the present experiments is very similar to that employed by Lesgold, Roth and Curtis (1979), who also varied topical continuity between anaphor and antecedent; to this extent, their results are open to the same interpretation as those presented here. One difference is that Lesgold et al. used definite noun phrase anaphora rather than pronouns; definite noun phrases are interesting because their inherent semantic content may be sufficient on its own both to disambiguate the reference and to act as a retrieval cue to relevant sections of memory. If this is the case, definite noun phrase anaphora might be expected to be less sensitive than pronouns to the amount and nature of the information given in the remainder of the target sentence. An interesting possibility is that the choice of the noun phrase used to refer to an individual may itself be a function of the expected ease of antecedent retrieval. In cases where antecedent retrieval is expected to be difficult, it may be most appropriate for writers to use a complex noun phrase, possibly with qualifying adjectives and relative clauses, in order to assist the reader in identifying the intended referent. Similar factors may affect the initial decision as to whether to refer to an entity by means of a definite noun phrase rather than a pronoun; Sanford and Garrod (1981) suggest that definite noun phrases are appropriate when the intended antecedent is no longer foregrounded.

The paper by Lesgold et al. (1979) is interesting because the authors do not in fact give a consistent interpretation of the foregrounding effects which they obtained. As described earlier, one of the important results which they present concerned the effects of adding an extra sentence at the end of the intervening backgrounding material; when that sentence referred to concepts linked to the antecedent of the subsequent definite noun phrase, reading times for the anaphoric

sentence were facilitated. In the first interpretation given for this effect, Lesgold et al. follow the assumptions of the working memory theory; they assert that "comprehension of the repetition sentences involves reinstatement of a portion of the memory representation of the passage, which then permits immediate matching of the target sentence's given segment." (Lesgold et al., 1979; page 299). This and other statements imply that they regard the repetition sentences as fully reactivating the antecedent concept itself, even though it is not explicitly mentioned. Later in the same paper, however, and without acknowledgement of the change in their theoretical position, Lesgold et al. describe the repetition effect in terms of its consequences for antecedent retrieval. It is argued, not that repetition of concepts linked to the antecedent actually reinstates the antecedent directly, but only that such repetitions facilitate reinstatement of the antecedent when it is referred to in the following sentence. The anaphoric noun phrase is seen as initiating a reinstatement search process which begins from currently active concepts in the short-term buffer carried over from the previous sentence; when these concepts are linked to the desired antecedent, as they are in the repetition conditions, location of the antecedent in the text representation will be facilitated. This model assumes, as has been assumed here, that information from the preceding sentence is active in working memory and can play a role in anaphora resolution.

Although Lesgold et al. present this interpretation in relation to the effect of repeating concepts linked to the antecedent after an intervening block of backgrounding sentences, it is evident that it can be applied equally well to the conditions in which all the intervening sentences dealt with topics related to the antecedent. The standard foregrounding effect is thus open to interpretation in terms of

antecedent retrieval processes. Furthermore, Lesgold et al. speculate that aspects of the hierarchical structure of the text may also influence the ease of antecedent retrieval; it is suggested, for example, that topical antecedents may be easier to retrieve because they are located at higher levels of the text representation.

The notion that anaphora resolution may involve the use of cues or guides to appropriate sections of the text representation is also very similar to the concept of the discourse pointer used by Carpenter and Just (1977a, 1977b). In Carpenter and Just (1977a), the discourse pointer is regarded as directly activating concepts or structures in the text representation; to this extent, the foregrounded information marked by the discourse pointer could be said to be in working memory. The notion that the discourse pointer actually causes the relevant concept to be activated is however omitted from Carpenter and Just (1977b), where the pointer is said simply to denote, designate or index the topic of discourse, which is seen as being the preferred location for integrating the current sentence with what has gone before. It is not being argued here that Carpenter and Just themselves regard the discourse pointer as a guide to sections of long-term memory rather than as a device which serves to directly activate particular concepts, but such an interpretation would evidently be consistent with the present suggestions. Kieras (1981) does assert that topic pointers are used to indicate memory representations for particular referents, and that such pointers are used to guide memory searches; however, his model assumes that pronoun antecedents are also held in short-term memory as well as the topic pointer.

The results presented by Cirilo (1981) may also be interpreted in terms of ease of antecedent retrieval. Cirilo did obtain the distance

effect on definite noun phrase anaphora, but on the basis of the present Experiment 3 it can be argued that this is explicable in terms of confounded topical discontinuities. More interesting, however, is the interpretation which Cirilo offers for the effect of antecedent height in the text hierarchy. Although this is first explained in terms consonant with a standard working memory theory, Cirilo then argues that higher-level information need not in fact be in the short-term buffer in order to be more accessible to anaphoric reference; instead, antecedent retrieval processes may be initiated from the highest levels of the text structure, and work downwards. This suggestion was also made by Lesgold et al. (1979), and would clearly result in more rapid location of high-level antecedents. This downward searching through the text representation need not be seen as an alternative to directed searching guided by target sentence content and cues from the preceding sentence, but may instead be a simultaneous process acting in parallel with directed searching.

Certain evidence presented by Sanford and Garrod (1981) can also be re-interpreted in terms of retrieval processes affected by text structure. In particular, the experiment described earlier, in which scenario dependent entities became more difficult to refer to following a time change taking the story action beyond the bounds of that scenario, can be readily accounted for. Such an account, rather than stressing the clearing of scenario-dependent entities from focus, would emphasise the role played by time changes as cues used to infer text structure, and as signals to guide macro-structure formation. It is striking that Kintsch (1977) emphasises the role of time changes as cues to transitions between story categories. In addition, the use of conventionalised knowledge in text comprehension is emphasised by van Dijk (1977a), who regards the information contained in stereotypical

representations of standard events or situations as playing an important role in macro-structure formation. It can therefore be argued that frame or scenario information may be used in assessing spatial and temporal discontinuities in the text, in order to determine whether or not they constitute a possible transition between different scenes; if the discontinuity indicates such a transition, readers are likely to initiate closure of the preceding structural category and begin construction of a new one, while macro-processes will operate to reduce and summarise the information contained in the old subsection of the text. In this way, large time changes will both increase the structural separation between characters in the old category and later pronouns referring to them, while the availability of that character will be reduced by macro-structure formation. These processing consequences of a stated time change would clearly render antecedent retrieval more difficult.

The above observations on various published effects on ease of anaphora resolution have thus shown how these effects may be accommodated within a retrieval theory of foregrounding. Such a theory seeks to identify the semantic and structural factors which guide and constrain antecedent retrieval processes, and is clearly general in application. In addition, it is evident that in some cases an explanation based on antecedent retrieval has already been offered for particular published results. However, it is not sufficient to show simply that a retrieval theory is capable of explaining the phenomena of foregrounding; reasons must also be offered for accepting such a theory in preference to a model based on working memory presence of antecedents. Theoretical observations were made earlier to suggest that a working memory theory would lack generality and require many ad hoc separations between different memory partitions and different strategies of pronoun assignment. Such theoretical criticisms offer one basis for rejecting the working memory model; yet it is obvious that such rejection cannot be made with confidence without empirical evidence to support it.

The principal difficulty in obtaining empirical evidence to separate the working memory and retrieval theories is that they make identical predictions concerning the effects of foregrounding. Each theory can readily account for the known influences on ease of pronoun interpretation and can therefore explain simple reading time differences between foregrounded and backgrounded conditions. Despite the fact that the two theories are based on very different processing models, they predict identical patterns of overall processing difficulty across experimental conditions. Fletcher (1981) goes so far as to assert that

the distinction between buffer models and models based on differential accessibility in long-term memory is simply a disagreement over the nature of short-term memory, rather than over its functions in comprehension. Clearly, if working memory models and retrieval models are functionally equivalent, evidence to distinguish between them may be impossible to obtain.

One research strategy which may surmount this difficulty would be to turn away from the investigation of influences on foregrounding and begin to consider the processes by which it operates. It was noted earlier that the study of suggested influences on ease of assignment, such as distance, topicality, or time changes, in fact simply amasses more instances of the phenomenon which is to be explained. A better approach would take a more analytic view of the alleged processes involved in the working memory theory, and by contrasting them with the suggested retrieval processes attempt to identify the different consequences which they may have. This approach would not involve looking for new influences on foregrounding, since any that were discovered could presumably be accommodated by both the competing theories; rather, it would involve giving a known foregrounding effect closer examination. Since the two theories make different assertions regarding the processing which accounts for foregrounding, it may prove possible to identify the features of a known foregrounding effect which are ultimately responsible for changes in ease of pronoun assignment.

Experiments 4 and 5 can in fact be re-interpreted in this light. The alleged assignment processes contained within the working memory theory require that foregrounded antecedents be held in a directly accessible short-term store; such working memory presence should facilitate recognition memory responses, but no evidence for this was

found. The retrieval theory, however, would not predict a foregrounding effect on recognition memory. This follows for two reasons. Firstly, the retrieval theory argues that more rapid accessing of antecedents in the foregrounded case is partly the result of an interaction between the sentence content which disambiguates the pronoun and the information in the intervening foregrounding material; the disambiguating target sentence content highlights a feature of the antecedent made salient by the intervening material, and acts as a retrieval guide to appropriate memory locations. This question is addressed in Experiments 6 and 7, but it is clear that since recognition tests do not provide such retrieval cues the foregrounding effect must at least be weakened. Secondly, it is noteworthy that while the working memory theory places foregrounded items in a general processing store, which must be as accessible to the processes attempting to make recognition decisions as to those attempting to resolve anaphoric references, the retrieval theory is based on cognitive operations which appear specific to discourse processing. The retrieval theory ascribes the foregrounding effect to processing strategies which are called into play during text comprehension; it is implausible that similar retrieval strategies are used in making conscious memory searches in response to a recognition test. Whatever strategies are used to carry out such a recognition task, they will clearly require use of the working memory system and would therefore be sensitive to current activation of the test word; but it is unlikely that these strategies will be similar to the assignment processes used in interpreting pronouns. If the retrieval theory is correct, therefore, antecedents would not be activated when foregrounded, and the conscious memory searches used in making recognition decisions would not be sensitive to the factors which influenced ease of pronoun assignment.

The results of Experiments 4-5 can therefore be seen as inconsistent with the working memory theory of pronoun assignment, but consistent with the retrieval theory. To this extent, the functional equivalence of working memory and retrieval theories, as asserted by Fletcher (1981), is not complete; with regard to recognition tests at least, they do make distinct predictions. It would, however, be preferable to identify separable predictions of the two theories with respect to ease of pronoun assignment itself; this is attempted in the following sets of experiments.

CHAPTER V

Pronoun Disambiguation and the Foregrounding Effect

- 1 Experiment 6
- 2 Experiment 7
- 3 General Discussion

Introduction

The most striking feature of the immediate mapping theory of pronoun assignment is its assertion that available disambiguating content is ignored. The first processing stage carried out when readers encounter a pronoun is claimed to involve attempted direct matches to working memory; in the foregrounded condition, the relevant antecedent is available there and matching succeeds. Only if the required antecedent is backgrounded is disambiguating semantic content used to elaborate the specification of the object of search and to guide retrieval of the antecedent from the long-term text representation. By contrast to this view, the retrieval theory of foregrounding argues that the disambiguating information in pronoun sentences plays a role in assignment even in the foregrounded case, both in locating the antecedent in memory and in verifying that a particular entity is in fact the correct referent. In addition, information from the immediately preceding sentence is claimed to be used as a pointer to substructures in the text representation, and so will also play a role in guiding antecedent retrieval.

It was suggested earlier that one source of the foregrounding effect could be found in the semantic relationship between the disambiguating content of the target sentence and the content of the sentences intervening between pronoun and antecedent. In the passages written for Experiment 3, this relationship was generally a close one in the foregrounded condition, but a more distant one in the backgrounded condition. As was suggested in the earlier description of the retrieval

theory, this difference could contribute to ease of antecedent retrieval in the foregrounded condition in a number of ways. Firstly, the specification of the required antecedent could be augmented still further by using information from the preceding sentence as well as from the target sentence itself; secondly, the defining attribute of the antecedent used to disambiguate the pronoun may have been made more salient by the foregrounding sentences, and hence less likely to have been deleted in the formation of semantic macro-structures; and thirdly, use of the preceding sentence as a pointer would direct the search process to an area of the text representation where the disambiguating content of the pronoun sentence could readily locate a match. Thus, at least some of the foregrounding effect could be due to interactions between the nature of the foregrounding manipulation and the immediate context of the pronoun.

The semantic relationship between target sentence and intervening material is thus a potentially important contributor to the foregrounding effect; however, this relationship arose as an almost incidental consequence of the manipulation being used here. In the experimental passages, it was necessary that the crucial antecedent be made clearly unambiguous in order to ensure correct assignment; this was in order to avoid the risk of subjects making incorrect mappings, which may for example have affected the results of Clark and Sengul (1979). Ambiguity was avoided by predicating something of the pronoun which was related to what had been said of the antecedent, but did not relate to what had been said of any of the other specific entities in the passage. Since the intervening foregrounding sentences were also written to deal with topics related to the antecedent, it necessarily followed that the predicate of the target sentence and the content of the foregrounding sentences were also related. Although the strength

and obviousness of this relationship, which will in future be referred to as predicate relatedness, clearly varied between different experimental passages, the connection was nonetheless present.

It is not, however, inevitably the case that the target predicate be related to the intervening foregrounding sentences. One way in which this could be avoided would be to deliberately define the crucial antecedent in two different ways. The antecedent entity could be described as having two distinct qualities, or being involved in two different events or activities. Whatever the nature of the two attributes, one of these could be used as the central topic of the foregrounding sentences while information relating to the other could be used to disambiguate the target sentence pronoun. This procedure would manipulate foregrounding precisely as before, since the intervening foregrounding sentences would continue to deal with topics related to the antecedent; and in addition, the pronoun would be disambiguated in the same way as before, since the target sentence predicate would be related to an attribute unique to the antecedent. However, as the attributes used for these two purposes would be different from each other, the target predicate would bear no direct relation to the intervening foregrounding sentences.

On a retrieval theory of foregrounding such lack of predicate relatedness would be expected to hinder pronoun assignment. Firstly, although both target sentence content and information from the preceding sentence would still be related to the antecedent in the foregrounded case, they could no longer be combined into a unitary augmented definition of that antecedent. To a certain extent, the two sources of information are likely to conflict. There may also be initial uncertainty as to whether the information from the preceding sentence

was in fact connected to the same individual as the target sentence. Secondly, and related to this, the target sentence content will not match to the substructure of the text representation indicated by the pointers carried over from the preceding sentence. Such a substructure is likely to be headed by information concerning the general topic area of the intervening sentences, and may thus seem an inappropriate location to search for an antecedent defined by the different information in the target sentence. These conflicts may lead to uncertainty as to where to begin the memory search, or regarding which attribute to adopt as a criterion for deciding when the antecedent has been found. Thirdly, since the attribute used in the target sentence has played no role in the interpretation of the intervening foregrounding sentences, it is likely to have been deleted during macro-structure formation; thus, in order to match target sentence content with the relevant antecedent attribute, retrieval of lower level text propositions will be required. Target sentences with unrelated predicates should therefore produce longer reading times than those with related predicates, owing to these hindrances to rapid antecedent retrieval. However, the foregrounded condition with an unrelated predicate should nonetheless not be equivalent to the standard backgrounded condition. This follows because of the continued availability of the crucial antecedent concept itself, even though the pronominally relevant attribute may be less accessible; in the backgrounded condition, the entire antecedent concept will be inaccessible. In addition, in the backgrounded condition the intervening sentences are even less related to the target predicate than the intervening sentences in the unrelated foregrounded condition. Thus, topical continuity with an unrelated predicate may be expected to produce some facilitation relative to the standard backgrounded condition, but less so than the foregrounded condition with a related

predicate.

In contrast to these arguments, the working memory theory of foregrounding predicts that predicate relatedness should have no effect on the strength of facilitation. This theory attributes the foregrounding effect to variations in the status of antecedents which are determined independently of whether and how they are referred to again. In principle, it is possible for a concept to be foregrounded and yet go unmentioned. Thus, provided that antecedent foregrounding has been successfully achieved before the pronoun is encountered, assignment should be easy and rapid however the pronoun is disambiguated. The working memory theory clearly excludes the possibility that the target predicate may play a role in assignment in the foregrounded condition.

This experiment therefore set out to test the effects of predicate relatedness on ease of pronoun interpretation. The experimental texts were constructed in the manner outlined above, as is described in detail in the Method section. If predicate relatedness was found to be important in obtaining the foregrounding effect, this would be considered more consistent with a form of retrieval theory than with the working memory theory. At the same time, the size of the effect of predicate relatedness was expected to give some indication of how much of the foregrounding effect could be explained in this way. Although it was suggested above that the unrelated foregrounded condition should not be as difficult as the standard backgrounded condition, an actual obtained equivalence between these two would suggest that other aspects of text structure than those involved in predicate relatedness made little contribution to foregrounding.

In addition to investigating the effect of predicate relatedness, this experiment also set out to confirm the existence of the basic foregrounding effect itself. This effect was demonstrated in Experiment 3, and again in Experiment 5 under slightly changed experimental conditions; in the present experiment, the replication was attempted using new experimental materials, in order to confirm the generality of the effect across passages. Although some researchers support the use of statistical methods which treat materials as a random effect in order to establish generality (Clark, 1973), the applicability of this approach to studies using specifically constructed texts is debatable; for such cases, replication on new materials is probably more appropriate. This approach is recommended by Wike and Church (1976). This experiment also set out to replicate the interaction of foregrounding with distance found in Experiment 3, but on this occasion only distances of up to four intervening sentences were used. This was considered sufficient to establish whether additional distance produced any additional effects on pronoun interpretation. The experiment also included reader skill as a factor, in order to give possible good and poor reader differences in sensitivity to foregrounding a final test.

One final change in this experiment was that questions interrogating the referent of the target pronoun were not included. These had twice failed to show any effect of foregrounding, while in Experiment 3 the significant effect of distance had only been seen in comparisons between distance 0 and distances 6 and 8. In addition, the question technique was criticised earlier as being at least in part a reflection of inferences made at the time of answering, rather than solely of ease of pronoun interpretation. Pronoun questions were therefore no longer considered likely to produce valuable data.

Subjects

Subjects in this experiment were twenty-four Plymouth Polytechnic students who participated either for a first-year psychology course credit or for a fee of £1.50. Subjects were divided into skill groups after the experiment was completed. Two further people who carried out the experiment were excluded from the analysis, and other subjects were run to replace them; one of these exclusions was a non-native speaker of English, while the other had claimed in the debriefing session to have been previously diagnosed as a dyslexic.

This second exclusion perhaps deserves a lengthier comment. Although reading skill was again of interest in this experiment, only individual variations within the normal range of adult ability were under consideration. Severely disabled readers, or those with a specific neurological impairment, were never seen as part of the population being sampled or to which the results were meant to be generalised. Because of this, despite the controversy over the meaning of the term dyslexia, and despite the fact that this subject did not elaborate on her completely unprompted claim, it was felt wisest to exclude her from the experiment. This decision was made without inspecting the subject's data.

Materials

Twenty-four new experimental passages were written for this experiment. Although some of these made use of a setting or group of

characters that had figured in one of the earlier experimental texts, the events being described here were new. The passages are given in full in Appendix 8.

The general structure of the passages was very similar to that used in previous experiments, with one important change. As before, the crucial antecedent noun occurred only once, in the final sentence of the constant introductory section; on this occasion, however, each antecedent entity was described as being involved in two actions, or possessing two attributes, rather than one. An example passage is given in Table 6.1.

In this passage, the introductory section describes members of an orchestra setting off on a tour. The crucial antecedent is the violinist, who is said to be both sitting in the front seat of the coach and reading a newspaper. These two attributes are logically distinct, in that assigning one of them to the violinist in no way requires the other also to be true; this distinctness was the case in all passages. In some passages the attributes were actions, in others they were more passive qualities, such as being grey-haired. The two attributes were always unique to the crucial antecedent within its own passage, and were not ascribed to any other character in that text. Either attribute could therefore be used to disambiguate the eventual pronominal reference.

Table 6.1

Experiment 6: Example Passage

The regional orchestra was going on tour again.
Once a year, the company travelled all over the region.
They gave concerts in town halls and schools, as well as theatres.
The coach arrived at their headquarters at eight in the morning.
All the players were ready, with their instruments and suitcases.
The flautist and cellist helped to stow the baggage.
The trombonist had brought a pack of cards to use on the journey.
The violinist took the front seat, and read a newspaper.

FOREGROUND BLOCK

This was really the best seat in the coach.
There was plenty of leg room, with no other seats in front.
It was nearest to the door, too, which was convenient.
Any trip is more pleasant in a comfortable seat.

BACKGROUND BLOCK

These tours were a welcome change to the usual routine.
Throughout the region, audiences were enthusiastic and appreciative.
The travelling was an important part of the orchestra's role.
Many places would never see live music otherwise.

TARGETS

He liked being able to watch the road ahead.
He began with the sports pages.

- 1) How often did the orchestra go on tour?
- 2) What time did the coach arrive?

Table 6.1 also gives the blocks of intervening material for this example passage. On this occasion, the blocks contained only four sentences each. As for previous experiments, the backgrounding block dealt with topics unrelated to the crucial antecedent; to meet the specific requirements of the present experiment, the sentences in the foregrounding block were written on a topic related to only one of the antecedent's defining attributes. This can be seen in the example passage, where the foregrounding block continues on the topic of sitting

in the front seat of the coach, with no mention of topics relating to reading the newspaper. In this way, although the foregrounding of the antecedent was manipulated exactly as before, there was always a second antecedent attribute with which the foregrounding material did not deal. It was considered possible that the order in which the two attributes were mentioned in the antecedent sentence might affect their salience, or the extent to which they were linked to the antecedent; for this reason, in half the passages the foregrounding block continued on topics related to the attribute mentioned first, while in half the passages it dealt with the attribute mentioned second. There was only one foregrounding block for each passage.

For each passage, there were two pronoun target sentences. The two target sentences for any one passage always contained the same pronoun, always in sentence-initial position. The pronouns used were : 'he' (13 times), 'she' (7), 'her' (1), and 'it' (3). In one of the target sentences, the pronoun was made unambiguous by material related to the attribute which was used in the foregrounding block; in the other sentence, the disambiguating material was related to the attribute which had not been mentioned in the foregrounding block. In both cases, the referent was clearly unambiguous. In one of the target sentences, therefore, the predicate was related to the foregrounding material, while in the other it was not. For each passage, the two target sentences were approximately equal in length.

There were three within-subjects factors in the experiment, with a total of twelve conditions. Each passage could be presented with either 0, 2 or 4 intervening sentences between the pronoun and antecedent; with the intervening material drawn from either the foregrounding or the backgrounding block; and with either the related or unrelated target.

As in previous experiments, with a gap of 0 between pronoun and antecedent the foregrounding factor was a randomly assigned dummy variable; it should also be noted that with intervening backgrounding material the target relatedness factor served merely to identify the two distinct sentences which could be presented. Relatedness was not expected to have any effect in the backgrounded condition.

An alternative method of passage construction would have involved writing two foregrounding blocks for each passage, one dealing with each of the two antecedent attributes. In this way, the possibility that the order of mention of the two attributes might be a significant variable could have been controlled, and more importantly the two target sentences could each have been presented in both relatedness conditions. This would have guarded against the possibility that one sentence may have been simply harder to read than the other. However, this method would have complicated the design and increased the number of passages presented to subjects; for this reason, given that the two sentences for each passage were similar in length and that there were twenty-four different sentences in each relatedness condition, it was decided that this extra safeguard was unnecessary.

Two concealed practice passages were selected for this experiment. These were based on two passages randomly selected from those used in previous experiments, but were slightly reduced in length. The same method was adopted when constructing the single passage that was used for explicit practice. The new versions of these three passages are given in full in Appendix 9. Only one of the two full-screen reading speed passages was used in this experiment; this was the first passage given in Appendix 4. All the passages, including the new experimental texts, had two comprehension questions. In the experimental texts, both

questions related to material in the constant introductory section; the referent of the crucial pronoun was not interrogated. No distractor passages were used, in order to reduce experiment running time; this was considered justifiable, given the wide variety in the structure of the experimental passages on this occasion.

Twenty-four different random passage sequences were constructed. Each of these began with the two concealed practice passages, always in the same order, and these were followed by a random ordering of the twenty-four experimental texts. A different random assignment of passages to conditions was used for each sequence, with two passages in each of the twelve conditions. Each of these distinct sequences was presented to one subject; the reading speed passage was always interpolated after the first concealed practice passage.

Procedure

Each subject attended a single one-hour session, in which a complete passage sequence was read. Passages were presented sentence by sentence, in exactly the manner described in previous experiments.

Results

Subjects were divided into skill groups on the basis of their reading speeds on the single full-screen passage; subjects with the twelve fastest reading speeds constituted the skilled reader group, while subjects with the twelve slowest speeds constituted the less-skilled reader group. It should be noted that the use of only one

reading speed passage on this occasion was justified by the high correlations found between the two passages in previous experiments.

Main Analyses

The objectives of this experiment fell into two categories; firstly, to attempt a replication on new materials of effects found in earlier experiments, and secondly to investigate the new issue of the role of predicate relatedness in these effects. The reading time data were analysed as a 2 x 2 x 2 x 3 design, with Skill, Foregrounding, Relatedness and Distance as factors; this single analysis encompassed both aspects of the experiment, but for convenience of presentation they will be reported separately. The replication results will be described first, followed by the results dealing with the effect of Relatedness. The summary ANOVA table is given in Appendix 19.

Table 6.2

Experiment 6: Mean Pronoun Sentence Reading Times (msec)

<u>Foregrounding x Distance</u>				
	<u>DISTANCE</u>			
	<u>0</u>	<u>2</u>	<u>4</u>	
<u>FOREGROUND</u>	2142	2493	2158	2264
<u>BACKGROUND</u>	2271	2906	2907	2695
	2207	2699	2532	

Table 6.2 presents the mean reading times in the foregrounded and backgrounded conditions, and at distances 0, 2 and 4. The main effect

of Skill was found to be significant, with $F(1,22)=4.36$, $p<0.05$, but this factor did not interact with any of the other experimental effects; skilled readers had a mean reading time approximately 700msec faster than less-skilled readers. It can be seen from Table 6.2 that the mean reading time in the foregrounded condition was 2264msec, while the mean reading time in the backgrounded condition was 2695msec; this difference was somewhat smaller than those obtained in earlier experiments but was nonetheless significant, with $F(1,22)=7.59$, $p<0.025$. The main effect of Distance was also significant, with $F(2,44)=4.41$, $p<0.025$; it can be seen that mean reading times were fastest at Distance 0, increased by approximately 500msec to distance 2, and remained at a relatively high level at distance 4. The Newman-Keuls procedure was used to test the differences between these three means. It was found that for the 0-2 comparison, $Q(3)=4.12$, $p<0.05$; for the 0-4 comparison, $Q(2)=2.72$, not significant; and for the 2-4 comparison, $Q(2)=1.41$, not significant ($df=44$ in each case). Thus, the main effect of Distance was due solely to the comparisons with distance 0; the critical value of Q for $p<0.05$ in the 0-4 comparison was 2.88, but the most important result was the near equality of distances 2 and 4.

These effects have therefore confirmed, on new passages, the effects found previously in Experiments 3 and 5. Pronoun sentences were read faster in the foregrounded condition than in the backgrounded condition, while in addition an increase in the amount of material intervening between pronoun and antecedent was seen not to lead to increased reading times. There was, however, one point of disagreement between this experiment and Experiment 3; although in Table 6.2 it can be seen that the effect of Distance was most notable in the backgrounded condition, with the distance 0 and distance 4 means being virtually equal in the foregrounded case, the Foregrounding x Distance interaction

did not reach significance, with $F(2,44)=2.13$. One possible explanation for this result concerns the means at distance 0. As in previous experiments, the Foregrounding factor at this distance was a randomly assigned dummy variable, and the 129msec difference between the foregrounded and backgrounded means must be assumed to be due to chance. However, the direction of this difference can be seen to be such as would increase any apparent effect of Distance in the foregrounded case, and reduce it in the backgrounded case. It is therefore possible that this random discrepancy may have obscured the Foregrounding x Distance interaction; the failure of this interaction to reach significance may be partly a consequence of lack of power in the analysis, rather than a reflection of genuine absence of such an interaction.

In order to make a further test of the Foregrounding x Distance interaction, a procedure similar to that described in Experiments 3 and 4 was carried out. For each subject, a single mean reading time was calculated over all eight passages read at distance 0; similarly, a single mean was calculated for each subject over the four passages read in each of the remaining four conditions. These means were then analysed as a 2 x 5 design, with Skill as one factor and with distance 0 as a whole being directly compared with the other genuine distance conditions. The ANOVA summary table is given in Appendix 19. The analysis produced a significant result, with $F(4,88)=4.36$, $p<0.005$, and the Newman-Keuls procedure was therefore used to test differences between the five conditions. It was found that distance 0 was significantly different from the two backgrounded conditions; for the 0-2B comparison, $Q(3)=4.02$, $p<0.05$, and for the 0-4B comparison, $Q(4)=4.02$, $p<0.05$ ($df=88$ in each case). The two backgrounded conditions did not differ from each other. Distance 0 was not, however, significantly different from either of the two foregrounded conditions;

for the 0-2F comparison, $Q(3)=1.64$, and for the 0-4F comparison, $Q(2)<1$. The two foregrounded conditions also did not differ from each other. From these results, it can be concluded that Distance was in fact having an effect only in the backgrounded case, but that this did not appear as a significant interaction in the original analysis owing to the discrepancy between the two means at distance 0.

It should be pointed out that, as in Experiments 3 and 4, the above reanalysis of the data was evidently not independent of the original analysis with distance 0 randomly divided into foregrounded and backgrounded conditions; by effectively analysing the data twice, the possibility of obtaining a significant result when no genuine effects were in fact present has been slightly increased. This possibility was of little consequence in Experiments 3 and 4, since the results of the two analyses were on those occasions identical. In the present experiment, however, the second analysis has produced significant effects which were not present in the original analysis. It is being argued here that this is due to the more accurate assessment of mean reading time at distance 0, achieved by calculating a single mean over four reading times rather than two separate means over only two reading times each; however, the possibility of Type I error must be noted.

The set of results described above has therefore successfully confirmed the findings of Experiments 3 and 5. The results concerned with the new topic of interest in this experiment, the relevance of predicate relatedness to the foregrounding effect, will now be described.

In the original analysis described above, it was found that the main effect of Relatedness was not significant, with $F<1$. For the

majority of the interaction effects involving the Relatedness factor, F values were also less than or equal to 1. The crucial result of interest concerned the interaction of Relatedness and Foregrounding, and only this effect proved to be significant, with $F(1,22)=9.91$, $p<0.005$. The mean reading times relating to this interaction are presented in Table 6.3. It can be seen that with the related target sentences, in which the predicate was related in content to the intervening foregrounding block, reading times in the foregrounded condition were approximately 700msec faster than in the backgrounded condition. With the unrelated target sentences, however, in which the predicate was not related in content to the foregrounding material, reading times in the foregrounded condition were only 150msec faster than in the backgrounded condition. It was found that the simple main effect of Foregrounding was strongly significant in the related condition, with $F(1,44)=15.29$, $p<0.001$, but did not even approach significance in the unrelated condition, with $F<1$. This result is fully compatible with the retrieval theory of pronoun interpretation, as it shows the foregrounding effect of topical continuity to be sensitive to target sentence content.

Table 6.3

Experiment 6: Mean Pronoun Sentence Reading Times (msec)

Foregrounding x Relatedness

	<u>PREDICATE TYPE</u>	
	<u>Related</u>	<u>Unrelated</u>
<u>FOREGROUND</u>	2184	2345
<u>BACKGROUND</u>	2892	2497
	2538	2421

One interesting feature of the results concerns differences in readability between the related and unrelated sentence sets. Although the main effect of Relatedness was not significant, this does not in fact imply that the two sets of sentences were equally easy to read. The simple main effect of Relatedness was found not to be significant in the foregrounded condition, with $F(1,44)=1.34$, but did prove significant in the backgrounded condition, with $F(1,44)=5.55$, $p<0.025$; as noted earlier, the backgrounded condition should have had no differential effects on the two sentence sets, and therefore this result implies that the unrelated set was in fact easier to read than the related set. This difference in the intrinsic readability of the two sentence sets is reflected in the 400msec advantage for unrelated sentences.

The consequences of this difference in the readability of the two sentence sets are as follows. Although the effect of Foregrounding was shown not to be significant in the unrelated condition, it is not possible to directly compare the related and unrelated sentences within

the foregrounded condition itself. This comparison is of some importance, as the results as they stand do not indicate whether unrelated sentences are still significantly different from related sentences within that condition. The possibility remains that unrelated sentences in the foregrounded condition, though not read significantly faster than in the standard backgrounded condition, may still not be read significantly slower than related sentences in the foregrounded condition. Such a result would imply that while the predicate content made some contribution to the foregrounding effect, it was not the sole explanation. This possibility cannot be assessed from the present data. If 400msec were added to the unrelated mean in the foregrounded condition, to compensate for the intrinsically faster speeds of those sentences, the resulting figure of 2745msec would be clearly greater than the mean for related sentences in the foregrounded condition; this difference, however, can obviously not be tested for significance. The above results have therefore demonstrated that predicate relatedness is an important factor contributing to the foregrounding effect, but do not allow an assessment to be made of the relative contributions of the predicate and the crucial pronoun itself.

Subsidiary Analyses

Having dealt with the questions of principal interest to the experiment, various other analyses will now be reported.

An analysis of comprehension accuracy was carried out on the two questions attached to each experimental passage; as in previous experiments, the analysis included the experimental variables as factors, but none of these were found to have significant effects. In

particular, for the main effect of Skill, $F < 1$; skilled readers had a mean accuracy of 3.667 (maximum=4), and less-skilled readers had a mean accuracy of 3.722. There was thus a high level of comprehension performance in the experiment, with the faster readers attaining the same standard as the slower readers.

The analysis of the intervening sentences was carried out much as in previous experiments. Firstly, a comparison was run between the foregrounded and backgrounded conditions over all four intervening sentence positions, using only the reading times from the eight passages presented at distance 4 as data; for each subject, the mean reading time for each position in the two foregrounding conditions was therefore calculated over four sentences. The analysis was a $2 \times 2 \times 4$ design, with Skill, Foregrounding, and Position as factors. The main effect of Skill did not reach full significance, with $F(1,22)=2.97$, $p < 0.1$, while for the main effect of Foregrounding, $F < 1$; the main effect of Position was, however, significant, with $F(3,66)=6.84$, $p < 0.001$. The source of this Position effect appeared to be the fact that reading times for the fourth intervening sentence were approximately 450msec faster than reading times for the first intervening sentence; such an effect was not found in previous analyses of the intervening material, but is clearly in line with the overall serial position effect previously found in the analyses of distractor passages. A second significant effect in the present analysis was the Foregrounding \times Position interaction, with $F(3,66)=3.38$, $p < 0.025$; it was notable that while the two foregrounding conditions had roughly equal reading times at the three latest positions, at the first sentence of the intervening material the reading time in the foregrounded condition was some 400msec slower than that in the backgrounded condition. This interaction, however, would not reach significance if conservative degrees of freedom were used.

The difference in reading time for the first sentences of the intervening blocks was revealed again in the second analysis of the intervening material, which looked at the first sentence position alone; this analysis employed twice as many data as the preceding analysis, since reading times from the distance 2 passages could also be included. The 2 x 2 analysis showed a significant effect of Skill, with $F(1,22)=5.75$, $p<0.05$, and also of Foregrounding, with $F(1,22)=6.21$, $p<0.025$. The effect of Skill was in the usual direction, with skilled readers being some 850msec faster than less-skilled, while the effect of Foregrounding was again due to the mean in the foregrounded condition being 400msec slower than in the backgrounded condition. Skill and Foregrounding did not interact, with $F<1$. This analysis therefore provides some support for the view that intervening foregrounding sentences are read with greater attention than intervening backgrounding sentences, and to that extent provides some confirmation of the result found in Experiment 3.

The final analyses to be described concern the consequences of including no distractor passages in the experiment. In previous experiments, it was usually the case that several subjects would comment in the debriefing session on the unusual or slightly disjointed structure of some of the passages; for example, in Experiment 5 eight subjects mentioned the sudden reversion to an old topic at the end of certain passages, with a final sentence that seemed unconnected to what had gone before. These subjects had clearly noticed the experimental structure. In the present experiment, however, fully twenty-three of the twenty-four subjects made some such comment in the debriefing, usually describing the structure of the experimental passages in some detail and also often mentioning that questions were only asked on the first few sentences. It is obvious that the absence of distractor

passages led to a much higher rate of reporting that the experimental structure had been noticed. Despite this, however, the desired experimental effects were still revealed. One possible explanation for this is that subjects' conscious knowledge and expectations of the passage structure had no effect on their reading strategies; a more plausible explanation is that subjects did not become fully aware of the passage structure until several examples had been read, and therefore their strategies did not change until later in the passage sequence.

As a test of this latter possibility, the data were reanalysed with the inclusion of a new factor, which will be termed the Sequence factor. Each subject read twenty-four experimental passages in a single sequence, and for the purposes of this analysis these sequences were divided into first and last halves. Because the orders of presentation had been randomised, it frequently happened that the two experimental passages in any one condition both occurred in the same sequence half for a particular subject; when this occurred, the mean of these two reading times was taken as the data point for that half, while the other sequence half was left with a missing plot. This procedure therefore resulted in a $2 \times 2 \times 2 \times 3$ table, with Sequence, Foregrounding, Relatedness and Distance as conditions. Cells in this table for each subject contained either actual pronoun sentence reading times, when the two passages in that experimental condition had fallen in different sequence halves; a mean of two reading times, when both passages in one condition had fallen in that half; or were left blank, when that sequence half had been without an experimental passage in the particular condition.

Owing to the large number of missing plots in this data matrix, it proved impossible to analyse as a single design. Instead, various

methods of collapsing the data into smaller designs were used. Firstly, the data were analysed as a $2 \times 2 \times 2$ design, with Skill, Sequence and Foregrounding as factors. Each of the four cells of this analysis was therefore based on six cells from the original data matrix; although some of these six cells were blank, no subject was without at least one filled cell to draw on. Data points in this 2×2 matrix were therefore the means of however many of the six cells were filled, for each subject. Table 6.4 presents the resulting overall means for the foregrounded and backgrounded conditions, in the first and last sequence halves; the ANOVA summary table is given in Appendix 19. The main effect of Foregrounding was highly significant, with $F(1,22)=11.99$, $p<0.005$; it will be noticed that the means for these two conditions were slightly altered by the averaging procedure. Also highly significant was the main effect of Sequence, with $F(1,22)=27.12$, $p<0.001$; it can be seen that reading times in the second sequence half were some 800msec faster than in the first sequence half. Finally, the Sequence \times Foregrounding interaction was also significant, with $F(1,22)=12.93$, $p<0.005$; it was most noticeable that while the foregrounding effect was strongly present in the first sequence half, it was greatly reduced in the second sequence half. The simple main effect of Foregrounding was found to be significant in the first half, with $F(1,44)=23.99$, $p<0.001$, but was not significant in the second half, with $F=1$. It therefore appears that subjects were in fact affected by their awareness of the passage structure, in that the desired experimental effect disappeared in the latter part of the presentation sequence. This is consistent with the comments made by several subjects, that they began to ignore the final sentence of the passages. The increase in speed in the second sequence half was evident in both the foregrounded and backgrounded conditions, though being much greater in the latter, and the simple main effect of Sequence was significant in both cases.

Skill showed no interactions with these effects.

Table 6.4

Experiment 6: Mean Pronoun Sentence Reading Times (msec)

Foregrounding x Sequence

	<u>SEQUENCE HALF</u>		
	<u>First</u>	<u>Last</u>	
<u>FOREGROUND</u>	2486	2001	2244
<u>BACKGROUND</u>	3247	2159	2703
	2866	2080	

The preceding analysis collapsed the data matrix over the Distance and Relatedness factors, in order to consider the effect of Sequence on Foregrounding. A second analysis was conducted with the data collapsed over the Foregrounding and Relatedness factors, in order to consider the effect of Sequence on Distance. Each of the cells in this 2 x 3 design was therefore based on four cells of the original matrix. The main effect of Sequence was again highly significant, but in this analysis the main effect of Distance did not reach full significance, with $F(2,44)=2.96$, $p<0.1$. This loss of significance was presumably due to the changed properties of the data, and the slightly altered means at each of the three distances. The Distance x Sequence interaction did not approach significance, with $F(2,44)=1.69$. A final analysis was run with the data matrix collapsed only over Distance; the cells in this 2 x 2 x 2 design, with Sequence, Foregrounding and Relatedness as factors, were thus based on three cells of the original matrix. This analysis produced significant main effects of Sequence and Foregrounding, and

repeated the significant interaction between them found in the first analysis; the Foregrounding x Relatedness interaction was also significant, but there was no three-way interaction with Sequence. The Sequence x Relatedness interaction was also not significant, with $F < 1$. The above analyses both included Skill as a factor, but this showed no interactions with the other effects.

It therefore appears that subjects' conscious recognition of the structure of the experimental passages did affect their performance, but only with regard to the Foregrounding factor. Since this effect took the form principally of a large reduction in reading times in the backgrounded condition, it may be concluded that subjects eventually ceased to undertake the extra processing involved in integrating the target sentences in that condition.

Discussion

Two important results were produced by this experiment. Firstly, the effects of foregrounding and of distance found in previous experiments were successfully replicated, and secondly the dependence of the foregrounding effect on predicate relatedness was clearly demonstrated. These two findings will be discussed in order.

The replication of the basic foregrounding effect is especially important on this occasion, as the experiment used new passages and therefore confirms the generality of the effect. Although the size of the effect was smaller in absolute terms than that found in Experiments 3 and 5, it was nonetheless strongly significant. The effect of

foregrounding as manipulated in these experiments can thus be accepted as genuine. Similarly, this experiment has confirmed that the number of sentences intervening between pronoun and antecedent is irrelevant to ease of pronoun interpretation, even in the backgrounded condition. Although the distances used here were shorter than those used previously, it nonetheless appears that antecedent availability is determined by topical continuity and text structure rather than by surface distance. Thus, the antecedent retrieval processes which even the working memory theory accepts must take place in the backgrounded condition must clearly be complex ones; such simple retrieval strategies as sequential backward searching could not account for the lack of an additional effect of increased distance.

It is interesting that reading times for the first intervening foregrounding sentence were slower than those for the first backgrounding sentence. In Experiment 3, it was found that overall reading times for foregrounding sentences were slower than for backgrounding, but for the first intervening sentence this difference was reversed. The reversal was not significant, and in Experiments 4 and 5 no effects whatever were found on the intervening material; yet it seemed clear that as the first backgrounding sentence constituted a topic change, while the first foregrounding sentence did not, reading times should be predicted to be slower for the backgrounding sentence. The present result is the precise opposite of this. One possible explanation for this is that the first foregrounding sentence does in fact constitute a form of topic change, and may involve more complex processing than the first backgrounding sentence. In the backgrounded condition, the intervening sentences bear no direct specific relation to the introductory block; thus, when the topic change is detected processes of closure and macro-structure formation may delete and reduce

all the preceding information indiscriminately, into a general statement of the setting of the passage. In the foregrounded condition, however, the first intervening sentence actively selects one of the many individuals mentioned previously, and gives more detail concerning that individual. This also constitutes a topic change from the preceding description of a set of individuals and their activities, with none of them given special emphasis, but on this occasion macro-processes must be selective in order not to delete specific information concerning the one individual who has now become the main character of the passage. Thus, while in the backgrounded condition reduction of the introductory section can proceed indiscriminately, in the foregrounded condition the process must be constrained. These differences in the reduction of information from the introductory block have been suggested here as one cause of greater antecedent availability in the foregrounded case, but it is also possible that they could account for slower reading times on the first foregrounding sentence. However, this interpretation must clearly be a cautious one; the different reading times for the first intervening sentence have appeared only in this experiment, and the effect is evidently not robust. It is possible that the effect on this occasion may simply be due to random unintended differences in the ease of reading the particular sentences in the different intervening blocks.

The demonstration of the effect of predicate relatedness was the crucial new outcome of the experiment. In the standard foregrounded condition, with a target sentence predicate which was related to the intervening sentences, reading times were significantly faster than in the backgrounded condition. When the target sentence had an unrelated predicate, however, reading times were no different from the backgrounded condition. It may also be noted that while the significant

main effect of foregrounding was reduced in absolute size compared to previous experiments, the size of the foregrounding effect with the standard related predicate was much closer to that obtained previously. The absence of a foregrounding effect in the unrelated condition clearly contradicts the working memory theory of immediate mapping, but is consistent with a retrieval theory stressing the importance of complex search and decision processes in pronoun assignment. In the related condition, topical continuity extends directly from the antecedent sentence to the target sentence itself, and provides retrieval paths and consistent antecedent specifications which will facilitate searches of the text representation. In the unrelated condition, continuity between the intervening sentences and the target sentence is breached, while the attribute used to disambiguate the pronoun has been reduced during construction of the text representation. Thus, retrieval would be expected to proceed more rapidly with a related predicate than with an unrelated predicate.

There is, however, one difficulty with the pattern of results which prevents a full interpretation of the relatedness effect. The design of the experiment was such that the target sentences in the related and unrelated conditions were always different, whereas the usual technique is to have constant target sentences in all conditions. One problem with this method is that it could be argued that the set of unrelated sentences was simply insensitive to the effect of foregrounding, for reasons unconnected with predicate relatedness; this, however, would be highly unlikely, especially as one conclusion from the present experiment is that the foregrounding effect generalises across materials. A more important danger, which was recognised in advance, was that the two sets of sentences might differ in their intrinsic readability, independent of context of occurrence. The fact that there

were twenty-four sentences in each set, and that the two sentences for each passage were of approximately equal length, was expected to avoid this danger; but as the difference between the related and unrelated sets in the backgrounded condition shows, the unrelated sentences were in fact read some 400msec faster.

This difference in intrinsic readability of the two sentence sets does not invalidate the conclusion that related target sentences showed a foregrounding effect while unrelated sentences did not, but it does prevent a possibly interesting comparison between the two sentence sets within the foregrounded condition itself. It is not possible to determine, within that condition alone, whether target sentences with unrelated predicates were read significantly slower than those with related predicates. If this difference were significant, with the unrelated case therefore being fully equivalent to the standard backgrounded condition, it would suggest that other factors than those involved in predicate relatedness made only a minimal contribution to the foregrounding effect. This would necessitate a careful and thorough analysis of the ways in which predicate relatedness may affect antecedent retrieval, with little attention being given to other possible influences. However, if the unrelated case proved not to be significantly slower than the related case, at the same time as not being significantly faster than the backgrounded condition, this would suggest that the processes involved in making use of predicate relatedness were not the sole factors contributing to the foregrounding effect. Thus, the importance of predicate relatedness to the foregrounding effect is not fully apparent from the present results. In order to attempt to resolve this question, as well as replicate the effect already obtained, a further experiment on this topic was carried out.

Introduction

The purpose of this experiment was to make a further test of the contribution of predicate relatedness to foregrounding, while correcting the design feature of Experiment 6 which had prevented comparisons between related and unrelated sentences in the foregrounded condition alone. In this way, it was hoped to discover whether predicate relatedness was sufficient in itself to account for the foregrounding effect, or whether other factors might contribute to the difference between the standard foregrounded and backgrounded conditions. The changes between this experiment and Experiment 6 are given in detail in the Method section.

One important change which must be mentioned here is that reader skill was no longer included as a factor. The reasons for this were fairly straightforward. Although reader skill was included as a factor throughout the preceding series of experiments, on only one occasion did it appear to interact with any other variable; and that single statistically significant interaction was itself rejected on closer investigation. It therefore appears evident that reading ability in these experiments bears no relation to the other effects of interest. The present investigation of skill differences in pronoun interpretation has clearly been unsuccessful; however, as noted in the discussion of Experiment 3 this lack of success may be attributable to the small range of abilities in the subjects used, rather than to genuine absence of such differences. Whatever the source of the present lack of skill interactions, it seems plain at this stage that further attempts to

investigate these possibilities would prove equally fruitless; and it was therefore decided to omit the skill factor from all future experiments. This decision led to one further methodological change. In previous experiments, it was impossible to counterbalance the presentation of materials sets since skill groups were only determined after the experiment was completed; in this and subsequent experiments, however, the balancing of materials sets across conditions was routinely carried out.

Method

Subjects

Subjects for this experiment were eighteen Plymouth Polytechnic students, one of whom participated for first-year psychology course credit and the remainder for a payment of £1.

Materials

There were six experimental passages on this occasion, which were randomly selected from the twenty-four used in Experiment 6. The chosen passages were numbers 5, 7, 9, 13, 19 and 21, as given in Appendix 8.

One major materials change from Experiment 6 was to ensure that both target sentences for each passage could appear in all experimental conditions, thus removing the danger that a specific sentence effect might obscure the desired experimental effect. To this end, a second foregrounding block was written for each passage, which dealt with topics related to the antecedent attribute which was not mentioned in

the existing block. Each passage in this experiment thus had three blocks of four sentences associated with it, any one of which could be inserted between the antecedent sentence and the target sentence. One of these blocks foregrounded the antecedent by reference to the first-mentioned defining attribute; one foregrounded the antecedent by reference to the second-mentioned defining attribute; while the third backgrounded the antecedent by dealing with topics unrelated to either attribute. The new foregrounding blocks for the six passages used in this experiment are given in Appendix 10. The pronoun target sentences were unchanged from Experiment 6, one being disambiguated by a predicate related to the first-mentioned attribute, and one by a predicate related to the second-mentioned attribute.

In this experiment, all four sentences of any one intervening block were always presented, so that pronoun-antecedent distance was not a factor. Each passage could thus be presented in six different forms, constructed from the three intervening blocks and the two target sentences. The crossing of these two variables is represented in Table 7.1, which schematically indicates the six possible forms of each passage; cell numbers in the table represent a pairing of the target and intervening block indicated in the row and column headings respectively. The subscripted numbers in the headings refer to the two antecedent attributes, mentioned first and second in the antecedent sentence; thus, for example, T1 indicates the target sentence disambiguated by a predicate related to the first-mentioned attribute.

Table 7.1

Experiment 7: Experimental Passages - Possible Forms

		<u>INTERVENING BLOCKS</u>		
		<u>FORE1</u>	<u>FORE2</u>	<u>BACK</u>
<u>TARGET</u> <u>SENTENCES</u>	<u>T1</u>	1	2	3
	<u>T2</u>	4	5	6

Although Table 7.1 implies a 3 x 2 design, the six text forms in fact constituted a single-factor, three-level design. It can be seen that in cells 1 and 5, where the row and column subscripts are identical, the disambiguating predicate of the target sentence is related to the foregrounding block; in cells 2 and 4, however, where the row and column subscripts do not match, the target predicate and the foregrounding material are unrelated. Cells 3 and 6 constitute the standard backgrounded condition. There were therefore two forms of each passage which fell into each of the three conditions. The three conditions will be referred to as the related, unrelated, and backgrounded conditions, where the first two are versions of the foregrounded condition. Each target sentence could thus appear once in each condition, while in addition each foregrounding block could also appear once in each relatedness condition.

Since on this occasion subjects were not divided into skill groups, it was possible to counterbalance the assignment of passages to

conditions. This was done using a Latin square.

In addition to the six experimental passages, there were two passages for concealed practice, and four for use as distractors. These passages were slightly shortened versions of texts used in previous experiments, and are given in full in Appendix 11. The inclusion of distractors in this experiment was another improvement on Experiment 6; these passages included questions on information occurring late in the text, and there were also passages which did not have a long list of characters near the beginning, in an attempt to conceal the patterns in the experimental materials. Appendix 11 also contains the one passage used for explicit practice. All passages in the experiment had two comprehension questions.

From these materials, eighteen passage sequences were constructed. All sequences began with the two concealed practice passages, always in the same order, and these were followed by a random ordering of the six experimental and four distractor passages. In each sequence, the assignment of passages to conditions was determined by the Latin square, but the eighteen sequences each had a different random passage ordering. Each passage sequence was presented to one subject. Subjects therefore saw all six experimental passages, one in each form and two in each condition; while each passage was presented eighteen times, three times in each form and six times in each condition.

Procedure

Subjects attended a single half-hour session, during which a complete passage sequence was read. The manner of presentation of the

passages was exactly as previously described, with subjects reading the explicit practice passage in the presence of the experimenter, and then being left alone to read through the sequence.

Results

Main Analysis

The primary data for this experiment were analysed as a single-factor, three-level design, comparing the related, unrelated and backgrounded conditions described above. As each subject read two passages in each condition, the analysis was conducted on the means of these two reading times; the ANOVA summary table is presented in Appendix 20.

Table 7.2

Experiment 7: Mean Pronoun Sentence Reading Times (msec)

<u>CONDITION</u>		
<u>Related</u>	<u>Unrelated</u>	<u>Backgrounded</u>
2471	2759	3187

Table 7.2 presents the mean pronoun sentence reading times in the three experimental conditions. The analysis produced a result which marginally failed to reach conventional significance levels, with $F(2,34)=3.23$, $p<0.1$; the critical value of F for $p<0.05$ was 3.28,

although with conservative degrees of freedom significance would be missed by a larger margin. It can be seen from Table 7.2 that the mean reading time in the related condition, which corresponded to the standard foregrounded condition of Experiments 3-5, was approximately 700msec faster than the mean reading time in the backgrounded condition; the mean reading time in the unrelated condition fell between these two extremes. Given that the difference between the related and backgrounded conditions was strongly predicted, and constituted a replication of an effect found in three previous experiments, it was decided that the obtained F value was sufficient to justify rejection of the null hypothesis. Comparisons between the three means were therefore made using the Newman-Keuls procedure. The related mean was found to be significantly faster than the backgrounded mean, with $Q(3)=3.57$, $p<0.05$; the unrelated mean, however, did not differ significantly from either the related or the backgrounded mean, with $Q(2)=1.44$ and $Q(2)=2.13$ respectively ($df=34$ for all three comparisons).

These results have therefore confirmed the findings of Experiment 6. Target sentences with related predicates showed a significant benefit from foregrounding, while sentences with unrelated predicates did not. In addition, however, this experiment has shown the unrelated condition to lie between the related and backgrounded conditions, being significantly different from neither.

Subsidiary Analyses

Three additional analyses of other aspects of the experiment were carried out. Firstly, comprehension accuracy on the two questions attached to each experimental passage was analysed, again as a

single-factor, three-level design. As expected, there was no effect of experimental condition, with $F(2,34)=2.03$; mean accuracy was 3.389, with a maximum possible of 4. Subjects therefore had an acceptably high level of comprehension performance.

The second additional analysis concerned reading times for the three blocks of intervening material. On this occasion, it was of interest not merely to compare reading times for foregrounding and backgrounding material, but also to test for possible differences between foregrounding sentences relating to the first and second attributes mentioned in the antecedent sentence. The analysis constituted a 3 x 4 design, with intervening Block and sentence Position as factors, and was conducted on the means of the two reading times for each block type for each subject. It should be noted that the counterbalancing procedure described earlier ensured that all intervening blocks were presented equally often. In the analysis, both the main effect of Block and the Block x Position interaction were found to be nonsignificant, with $F<1$; readings times for the three types of intervening material therefore did not differ. The main effect of Position did prove significant, however, with $F(3,51)=5.62$, $p<0.025$. It was most noticeable that mean reading time for the first sentence position was 3193msec, while mean reading time for the fourth sentence position was 2575msec; subjects therefore read faster on the later sentence positions.

This effect was confirmed in the third additional analysis, which considered the effect of sentence serial position in the four distractor passages. As the distractor passages had been shortened for this experiment, the minimum length of passage was only ten sentences; the analysis therefore constituted a single-factor, ten-level design, and

was conducted on the means of the four reading times at each serial position for each subject. Extra sentences in passages longer than the minimum were ignored. The analysis produced a significant result, with $F(9,153)=4.35$, $p<0.05$, although this would miss full significance with conservative degrees of freedom. On this occasion, it was found that the first sentence did not have an exceptionally long reading time, with positions one and two both yielding mean reading times of about six seconds; there was, however, the usual steady fall in reading times over the following serial positions up to position six, when reading times stabilised at approximately three and a half seconds.

Discussion

The principal purpose of this experiment was to test the relative position of the unrelated foregrounded condition compared to the standard foregrounded and backgrounded conditions. The results indicated that the unrelated condition lay between the two other conditions, being significantly different from neither; the standard foregrounded and backgrounded conditions themselves were significantly different, however, so replicating the basic foregrounding effect.

It therefore appears that topical continuity between the antecedent sentence and the following sentences is not sufficient on its own to ensure easier pronoun interpretation than in the backgrounded case. A significant foregrounding effect is obtained only when the predicate or disambiguating content of the target sentence is related to the topic of the intervening sentences. At the same time, however, topical continuity through the intervening material evidently makes some

contribution to the foregrounding effect, since the unrelated condition was not found to be significantly slower than the standard foregrounded condition. In this experiment, in fact, the unrelated condition fell somewhat nearer to the related condition than to the backgrounded condition. It must therefore be concluded that while predicate relatedness is an important factor in obtaining the foregrounding effect, as was shown in both this experiment and Experiment 6, nonetheless other influences must also be at work.

It is interesting to note that this experiment did not simply replicate the effects of Experiment 6 on new subjects, but also extended them to partly new materials. All of the passages used here included newly written foregrounding blocks, while in addition the target sentences which had previously entered only into one or other of the two relatedness conditions now entered into both. Clearly, the use of wholly new materials would be a more convincing demonstration of generality, but these changes do add to the strength of the evidence for the obtained effects. It is also interesting that the absolute size of the foregrounding effect in this experiment was very similar to that obtained in Experiment 6; considering only the related sentences in that experiment, the standard foregrounded condition was approximately 700msec faster than the backgrounded condition, while the corresponding difference in this experiment was also 700msec. This estimate of the size of the foregrounding effect is clearly more informative than the main effect of foregrounding in Experiment 6, which was attenuated by the lack of effect on unrelated sentences.

A final instructive comparison with Experiment 6 is to be found in the debriefing comments made by subjects. In the earlier experiment, almost all subjects reported having noticed the unusual structure of the

passages, and presumably as a result of this the foregrounding effect failed to appear in the latter half of the experimental sequences. This experiment included distractor passages in an attempt to prevent subjects becoming aware of the experimental manipulation, while the reduced number of passages read by subjects should also have made it more difficult to detect the constrained passage structure. As it turned out, only nine of the eighteen subjects in this experiment reported having noticed the structure. This is an improvement over the previous experiment, and thus to some extent confirms the value of including distractor passages among the experimental materials; how much of the improvement was simply due to the shortened experimental running time cannot be determined.

These experiments were concerned with the role played in pronoun assignment by the disambiguating information which may be predicated of the pronoun. If in the foregrounded case pronouns are mapped directly onto antecedents that are immediately available in working memory, the nature of the remaining pronoun sentence content would have to be considered irrelevant to the assignment process; this prediction should be especially strong in the case of sentence-initial pronouns such as those used here, since these pronouns would be assigned before the remainder of the sentence was processed. The results presented here, however, showed that the nature of the pronoun sentence content does in fact play an important role in determining the ease of pronoun assignment. Pronoun antecedents were foregrounded in these experiments in the manner found to be effective in Experiments 3 and 5, in which the crucial variable was seen to be topical continuity through the intervening material; here, however, this manipulation was found to produce significant facilitation of pronoun sentence reading times only if the predicate of the target sentence itself was also related to the intervening sentences. Pronouns referring to the allegedly foregrounded antecedent were no easier to interpret than in the backgrounded condition, if the information predicated of them bore no relation to the foregrounding sentences. The fact that in Experiment 7 the unrelated condition was also found not to be significantly slower than the related condition slightly confuses the findings, but does not invalidate the conclusion that unrelated sentences show no significant benefit from topical continuity. To this extent, it is evident that to describe antecedents as being foregrounded at a point where they are not in fact

being mentioned is misleading. Foregrounding is not an independently determined quality of the status of antecedents, but is determinable only relative to a particular anaphor in a particular context of use.

This general conclusion regarding the nature of foregrounding is consistent not only with the present results showing the effects of the content of the target sentences, but also with the results of Experiments 4 and 5 which indicated no differences in the degree of activation of items that were known to be foregrounded or backgrounded relative to pronominal reference. The conclusion is also consistent with the various findings in the literature which demonstrated the importance of syntactic and semantic relationships between pronoun and antecedent in determining the preferred pronoun assignment. It is being argued here that these patterns of effects, and specifically for present purposes the effect of predicate relatedness, can be treated most satisfactorily within a theory emphasising the retrieval of antecedents by means of complex search and decision processes. The effect of predicate relatedness is in apparent contradiction of the working memory theory of direct, mapping, but can be readily explained in terms of differences in the ease of antecedent retrieval.

An effect of predicate relatedness on pronoun assignment may possibly be accommodated within the working memory theory of foregrounding, but only by abandoning the notion of immediate mapping based on a highly restricted specification of the required antecedent. So far, little consideration has been given here to the nature of the antecedent representation which is assumed to be held in working memory; however, it is clear that this must consist of more than simply the syntactic specification which would match the characteristics of a following anaphoric pronoun. In the Kintsch model, for example, the

short-term buffer is assumed to hold sets of complete propositions, while in the model put forward by Sanford and Garrod (1981) working memory nodes representing individual entities are seen as part of a semantic network giving details of such things as their role in current scenes or events. Since this information is assumed to be present in working memory, if pronoun assignment were seen as involving attempts to locate an entity in working memory which matched in full the extended specification based on target sentence content, the effect of predicate relatedness could be explained. Given that the crucial antecedent was defined in two ways, it is likely that only the attribute which continued to be relevant to the foregrounding sentences would be held as part of its representation in working memory; thus, in the related case an extended definition of the required antecedent would find a match in working memory, while in the unrelated case it would not. The reading time difference would be explained by the long-term memory access which would be required to match the extended definition of the pronoun in the latter case.

This account of the effect of predicate relatedness is therefore based on the logical separation, noted earlier, between working memory presence of foregrounded antecedents and direct mapping of pronouns. The relatedness effect can be explained if mappings to working memory are seen as using more than just the specification derived from the pronoun itself. However, relaxation of the direct mapping component of the working memory theory reduces its value considerably. The main virtue of antecedent presence in working memory is that it limits the search domain and allows unambiguous assignment; this advantage is eroded if the search process makes use of augmented definitions of the required antecedent. Such extended definitions could equally well identify a unique referent if they were allowed to operate directly on

the long-term text representation.

In order to account for the obtained effect of predicate relatedness on pronoun sentence reading times within a working memory theory, at the same time as preserving the important notion of direct mapping, it would be necessary to attribute the effect to integrative processes taking place after immediate assignment had occurred. There are two ways in which such processes could produce longer reading times in the unrelated case. Firstly, even under the working memory theory readers must have some method of detecting and recovering from incorrect initial assignments; this may be necessary when reading a badly written text, for example, or when for some other reason the immediate mapping proves erroneous. It might therefore be the case that the disambiguating content is compared against the full working memory representation of the antecedent, in an attempt to verify the assignment; conflict between the two pieces of information in the unrelated case would lead to corresponding processing difficulties. It is clear, however, that this suggestion of post-assignment verification of a mapping is fully equivalent to use of augmented information in the initial mapping itself, and in a similar way erodes the main advantage of the working memory theory. The second possibility is that, even if the initial assignment were not deliberately double-checked, the inability to unify the incoming information from the target sentence with the working memory representation of the selected referent would require that that representation be substantially changed. In the unrelated case, therefore, long-term memory access might be required to locate information relevant to the newly reintroduced antecedent attribute, to which the new information could be connected; while in the related case, integration could proceed in a straightforward manner.

This latter suggestion is perhaps the most plausible way of accommodating the relatedness effect to a working memory theory of foregrounding. The suggestion attributes the effect not to processes of pronoun assignment, but to processes of integrating predicated information with the representation of an individual, while the retrieval theory attributes the effect to the assignment processes themselves. Evidently, there is no way in which the present results can discriminate between these two accounts. Data relevant to the problem will be presented in the following experiments, but for the present the uncertainty will be left on one side.

Before continuing, it should be pointed out that there is one possible explanation of the present results which is quite unconnected with pronoun interpretation processes. Rather than attributing the faster target sentence reading times in the related foregrounded case to higher-level features of text comprehension and integration, it may be possible to explain them on the basis of lower-level processes of word recognition. Much research on word recognition has shown that the presence of a semantically related context can lead to significant increases in the speed of identification; for example, research reviewed by Stanovich (1981) showed that subjects could identify words significantly faster if they were semantically congruent with a sentence fragment presented immediately before. It is therefore possible that semantic connections between the foregrounding sentences and the target sentence predicate led to facilitation of word recognition in the related case; overall sentence reading times would be slower in the unrelated case because such facilitation was absent. However, there are reasons for doubting whether contextual effects on word recognition would be sufficient to account for these results. Firstly, the research on such context effects has been concerned primarily with effects

occurring within single sentences; a common technique is to present the final word of a sentence separately after presentation of the earlier part of the sentence, and vary whether that word is or is not consistent with the preceding fragment. Whether similar context effects could occur across successive sentences, as would be necessary to explain the present results, is unclear. Secondly, the size of the context effects presented in the literature is frequently quite small, and would be insufficient to account for the 300msec reading time difference found here between related and unrelated conditions; this is especially likely to be the case, given that the closeness of the semantic relationship between target content and intervening sentences was not as great as is usual in context experiments. Thirdly, Mitchell and Green (1978), in an explicit test of context effects in text processing, found little evidence for such facilitation of word recognition. Hence, although the possibility of context effects during continuous reading is one which has received relatively little investigation, it appears that such processes are not a likely cause of the effect found here.

It therefore appears that the effect of predicate relatedness on pronoun interpretation is genuine, and the explanation suggested here is that the effect is based on complex antecedent retrieval processes. Suggestions concerning how relatedness might influence these processes were made earlier. The present experiments, however, were designed simply to demonstrate the existence of the overall relatedness effect itself, and not in order to analyse or test the processes by which it came about; the principal significance of the results is that they are more easily accommodated within a retrieval theory of foregrounding than a working memory theory. However, it is clear that this demonstration of the existence of the effect requires that experiments be conducted with a more narrow purpose, designed to validate and extend the

suggested ways in which the effect occurs. There are several avenues, within the context of the retrieval theory, which such research might explore. One of the suggestions made, for example, was that information from the target sentence and from preceding sentences could be combined in the related case into a unified specification of the required antecedent; the assistance which this would give to retrieval processes is a special case of the more general prediction that antecedent location should be easier the more defining information there is available to guide the search. Experiments might therefore be conducted to investigate the effect of varying the amount of information which is predicated of the target pronoun; assignment should prove to be easier, the more such information there is.

One interesting possibility raised by this is that target sentence reading times in the standard backgrounded condition may also be sensitive to the effects of predicate content. Since retrieval processes are accepted to occur in the backgrounded case even under the working memory theory, this would be an informative research topic in any event. Pronoun sentences occurring after a backgrounding topical discontinuity will presumably be easier to interpret if they contain greater amounts of information which may be used as a retrieval cue. A question related to these effects of the amount of available defining information concerns possible effects of the degree of closeness between that information and the original qualities or actions which were attributed to the antecedent. The present experiments used a coarse and unquantified manipulation of relatedness, but more subtle and controlled alterations in this variable might be expected to produce a continuum of target sentence reading times. Pronoun sentences would presumably be easiest to interpret if the disambiguating content exactly repeated, either verbatim or in paraphrase, the original statement of the

antecedent attribute.

Another suggested factor contributing to the relatedness effect concerned the information contained in higher-level macro-propositions. With intervening foregrounding sentences that dealt with only one attribute of the antecedent, information concerning the other attribute would not be included in superordinate levels of the macro-structure, and hence that attribute would be both less effective as a retrieval cue, and less available to disambiguate the target pronoun. This suggestion regarding macro-structure formation could be investigated using memory tests. It is known that free recall of texts is more likely to include an item of information if that information was entered into the macro-structure than if it was not, especially if recall is delayed (Kintsch, 1977); hence, assuming texts with the type of structure used here to be presented to subjects without the final target sentence, recall protocols would be more likely to mention the attribute which figured in the foregrounding sentences than the one which did not. In the backgrounded condition, the two attributes would be recalled approximately equally often, but the recall level would be very low as neither piece of information would have been preserved during macro-structure formation.

There are thus a number of ways in which the processes underlying the effect of predicate relatedness could be further investigated. However, one important conclusion from these experiments is that predicate relatedness alone is not sufficient to account for the foregrounding effect; this was shown by the intermediate positioning of reading times in the unrelated condition of Experiment 7, which were significantly different from neither the related foregrounded condition nor the standard backgrounded condition. The nature of the other

factors which may be at work must therefore also be considered. Rather than there being additional distinct processes involved, however, it is probable that the processes which lead to the relatedness effect within the foregrounded condition are also responsible for the further effect of complete backgrounding. For example, if in the unrelated condition target sentence content and information from the preceding sentence cannot be combined into a unified specification of the required antecedent, this is likely to be still more the case in the backgrounded condition. Again, if information in the preceding sentence is an apparently unhelpful pointer to the long-term text representation in the unrelated case, as it disagrees with the antecedent specification derived from the target sentence, in the backgrounded case the available pointer will genuinely have no value, since it has no connection with the antecedent and therefore will not lead back to the appropriate section of memory. Another example of this more extreme operation in the backgrounded condition of processes accounting for the relatedness effect has already been mentioned; macro-structure formation in the backgrounded case will lead to complete loss of specific antecedent information from the introductory sentences, while in the foregrounded condition only one or other of the antecedent attributes will be deleted.

Further application of the processes accounting for the relatedness effect may thus lead to the even longer reading times in the backgrounded condition. However, there is at least one factor which may operate in the backgrounded case but not in the unrelated foregrounded case. This is the possibility, described in the original presentation of the retrieval theory, that backgrounding may lead to greater structural separation between the pronoun sentence and the antecedent sentence. If additions to the text representation based on intervening

foregrounding sentences can correctly be seen as the forming of branching structures extending downward from the antecedent node, it is clear that this must occur in the unrelated condition just as much as in the related condition; only in the backgrounded case will new additions to the text representation form a distinct branching structure independent of that which holds the antecedent. Only in the backgrounded case, therefore, will antecedent retrieval be hampered by the fact that the required referent is in a different substructure of the text.

One interesting possibility is that antecedent retrieval in the backgrounded case may be facilitated by explicit signals of structural changes. Kintsch (1977) and van Dijk (1977a) suggest that text connectives such as 'however...', 'but...' and 'so...' are used by readers in identifying transitions between story categories, and serve to cue macro-structure formation. By the same token, it may be that such expressions as 'meanwhile...', 'to return to...' or 'as was said earlier...' act as indicators of a reversion to an old topic, or the re-opening of a prior text substructure. If such expressions were used prior to the anaphoric reference in the backgrounded case, retrieval processes may be facilitated owing to this warning that they would be required, and to the signal that the intended antecedent was likely to be in a structurally separate category.

Suggestions concerning the ways in which predicate relatedness and backgrounding may produce their effects on antecedent retrieval will now be left aside. It is evident, however, that there are many ways in which such effects could occur, and that the retrieval theory has thus generated several new suggestions for research.

CHAPTER VI

Processing Consequences of Incorrect Foregrounding

- 1 Experiment 8
- 2 Experiment 9
- 3 General Discussion

Introduction

In the discussion of the predicate relatedness effect obtained in Experiments 6 and 7, it was pointed out that this effect could be accommodated within a working memory theory of pronoun assignment if the assumption of immediate mapping was given up. It was argued, however, that direct assignment processes based on limited information were a central component of the working memory theory, and could not be given up without seriously reducing the value of such a theory. The principal advantage of working memory presence of antecedents is that it allows direct matching of pronouns without the use of semantic or structural disambiguation; allowing such inferences to play a role in pronoun assignment would make it possible to explain foregrounding phenomena more parsimoniously in terms of a retrieval theory. It is therefore evident that if immediate mapping is found not to take place in assignment of pronouns to foregrounded entities, the basis of the working memory theory will be weakened. This experiment set out to test one prediction of the immediate mapping component of the working memory theory, and to contrast it with predictions derived from the retrieval theory.

The feasibility of the immediate mapping strategy is dependent on the fact that texts are normally cooperatively structured such that pronouns do refer to whatever entity is currently foregrounded. If the intended antecedent were not foregrounded, it would be considered inappropriate to refer to it by means of a pronoun. Thus, a strategy which maps pronouns directly onto any entity in working memory which is

acceptable on such grounds as gender and number will usually be successful. However, this strategy will lead to significant processing difficulties if the entity in working memory is not in fact the intended antecedent. This might be the case in texts such as those used here, for example, if the sentences intervening between pronoun and antecedent dealt with topics related to some other individual from the introductory block than the actual referent of the pronoun. In such circumstances, the initial assignment of the pronoun to the foregrounded entity will be incorrect, with this only becoming apparent when sufficient disambiguating content has been processed and found to be incompatible with the selected referent. Discovering the erroneous assignment will lead to recovery procedures designed firstly to cancel that assignment and erase any memorial structures it may have produced, and secondly to search the text representation for the intended antecedent. Clearly, with this additional processing to be carried out, reading times for pronoun sentences should be longer when the wrong entity is foregrounded than when the correct entity is foregrounded.

The most interesting prediction of the working memory theory, however, and the one most relevant for discriminating it from the retrieval theory, concerns the comparison between the wrong foregrounded condition and the standard backgrounded condition. In the backgrounded condition which has been used here, the sentences intervening between pronoun and antecedent deal in general terms with aspects of the passage setting or events which are not specifically linked to one individual. As such, they ensure that all unique entities are backgrounded to the same extent. Thus, when a sentence-initial pronoun is encountered, the direct mapping strategy will fail at once, since there is no acceptable antecedent in working memory. In this way, it will be immediately apparent to the reader that comprehension of the pronoun will require

the use of retrieval processes based on disambiguating sentence content; and these processes will be initiated as soon as the rest of the sentence, or part of the sentence, has been read. It is therefore evident that the standard backgrounded condition should produce less processing difficulty than the wrong foregrounded condition. In the wrong foregrounded condition, the reader must detect the initial erroneous assignment, abort any integrative processes which have been begun on the basis of it, and re-assess the sentence content as an antecedent retrieval cue rather than as simple additional information concerning a known entity. The decision to execute retrieval processes will thus be delayed, while the ability to make use of the sentence content to guide those processes may be reduced owing to initially inappropriate assessments. In the standard backgrounded condition, however, retrieval processes will be carried out promptly, while the sentence content will be processed with the deliberate goal of extracting antecedent retrieval cues. These differences in processing between the two conditions should therefore lead to a clear reading time advantage for the standard backgrounded case.

The working memory theory thus predicts that pronoun sentence reading times when the wrong entity is foregrounded should be longer than in both the standard foregrounded condition and the standard backgrounded condition. The retrieval theory, however, predicts only that such reading times should be longer than in the foregrounded condition. If pronoun assignment depends on antecedent retrieval processes based on available disambiguating content, the sentence-initial pronoun will not in fact be assigned to the incorrect entity in the wrong foregrounded case. Rather, processing in the wrong foregrounded condition should be identical to that in the backgrounded condition; in both cases, information carried over from the preceding

sentence will be inconsistent with the target sentence and will in any event be unrelated to the memory location containing the antecedent, while the representation of the antecedent itself is likely to have been reduced during macro-structure formation and to be in a structurally separate branch of the text hierarchy. Since retrieval processes will be subject to identical constraints in the two conditions, and since on the assumptions of the retrieval theory these processes will be initiated equally readily in either case, it can be predicted that pronoun sentence reading times in the two conditions should be equal. Both conditions should also be slower than in the standard foregrounded condition.

The prediction that a wrong foregrounded condition should produce slower reading time than a neutral backgrounded condition was among the questions studied by Frederiksen (1981). Frederiksen presented subjects with short passages one sentence at a time; each passage contained an antecedent sentence which mentioned two entities, one of which was referred to by a pronoun in the target sentence. One additional sentence intervened between pronoun and antecedent, and either referred to the non-antecedent noun from the first sentence or was more neutral in content. The intervening reference to the non-antecedent noun was seen as causing that entity to be foregrounded. The nature of the more neutral intervening sentences was not specified, but presumably these may be comparable to the standard backgrounded condition used here. The results showed that target sentence reading times were significantly slower when an incorrect referent was foregrounded than in the neutral condition. Frederiksen interpreted this result in terms of the reduction in the topical status of the crucial antecedent, with consequent increases in the time needed to locate it when processing the target sentence.

Frederiksen's experiment is an interesting expression of the hypothesis that wrong foregrounding should prove a more difficult problem for pronoun interpretation than neutral backgrounding, but his results cannot be considered an adequate demonstration of this effect. The principal reason for this is that no comparison is presented between the condition with a neutral intervening sentence and one in which the intervening sentence foregrounded the crucial antecedent without direct reference to it. Although the neutral condition produced longer reading times than a condition in which the intervening sentence explicitly referred to the crucial antecedent, this is clearly not a suitable comparison. In addition, subjects were not simply reading the passages for comprehension, but were making overt judgments concerning the referents of the crucial pronouns; the task demands involved in making such judgments may have affected the pattern of results.

The present experiment therefore made a further test of the processing consequences of foregrounding an entity which was not in fact the referent of the crucial target pronoun. This was done using a text manipulation very similar to that used throughout these experiments; intervening sentences were written which dealt with topics unrelated to the intended antecedent but closely related to one of the alternative individuals mentioned in the introductory block. Further details are given in the Method section. The important question concerned whether or not the wrong foregrounded condition would produce longer reading times than the backgrounded condition, as predicted by the working memory theory.

Method

Subjects

Subjects for this experiment were eighteen Plymouth Polytechnic students, who were paid for their participation.

Materials

Six experimental passages were written for this experiment. Some of these were loosely based on scenes and character groups that had been used in previous experiments, but the bulk of the passage content was new. The passages are given in full in Appendix 12. The experiment also used one explicit practice, two concealed practice, and four distractor passages; these were exactly the same texts as were used for those purposes in Experiment 7. There were two comprehension questions on all passages used in the experiment.

An example of the experimental passages is given in Table 8.1. As in all previous experiments, the introductory section here sets the scene and presents several characters to act as alternative choices for the eventual target pronoun. The essential feature of the passages on this occasion was the presence of two crucial antecedent entities rather than one; these two entities were always mentioned together in the last sentence of the introductory section, and were always equivalent in gender and number. Thus, in the example passage the two crucial antecedents are Carol and Pat. For each passage, there were two alternative target sentences, either of which could be presented; one of these target sentences referred anaphorically to the first-mentioned

Table 8.1

Experiment 8: Example Passage

On Saturday night, a group of friends met in the pub. They had virtually grown up together, living in one neighbourhood. Although still friends, they had all gone their separate ways. Gary had joined the army, and was home on leave. Theresa was a student, and Clive worked as a bank clerk. Carol was a qualified nurse, and Pat was a typist in an insurance firm.

FOREGROUND 1

Working in the health service was a demanding profession. Caring for the sick was hard work, and emotionally draining. The hours were long, and the patients needed constant attention. It took total dedication to be able to stand the pace.

FOREGROUND 2

It was very boring to spend all day typing. Some of the documents were complex, with long and technical words. Sitting at the typewriter seemed to put a strain on the back. Trying to copy from poor handwriting made matters even worse.

BACKGROUND

It was a nice little pub, with a very pleasant atmosphere. In the corner, a small open fire was burning. The ceiling was criss-crossed with dark oak beams. The tables and chairs were also of old, sturdy wood.

TARGETS

She was working on a long-stay surgical ward.
She especially hated doing long business letters.

- 1) Why were the group of people such good friends?
- 2) What day of the week was it?

antecedent while the other referred to the second-mentioned antecedent. Because of the syntactic equivalence of the two antecedents, the same pronoun could be used in both targets of any one passage. The correct referent of the pronoun was made clear in the usual way, by means of a disambiguating predicate of the target sentence which was related to the defining attribute of one or other of the antecedents. Thus, the first target sentence given in the example passage must refer to Carol, while the second must refer to Pat. The pronouns used in the experimental

passages were 'he' (3 times) and 'she' (also 3 times); the pronouns always appeared in sentence-initial position.

For each passage, three blocks of intervening material were written. These were always four sentences long, and all four sentences of any one block were always presented; distance was not a factor in the experiment. Of the three blocks, one was written with the intention of foregrounding the first-mentioned crucial antecedent; as in previous experiments, this was done by writing sentences which were related to the defining attribute of that antecedent. Thus, in the example passage, the first block of intervening sentences acts to foreground Carol, by dealing with the topic of nursing. The second block of intervening material was intended to foreground the other crucial antecedent; in the example passage, the second intervening block acts to foreground Pat, by dealing with topics related to typing. It can be seen that the material in these two blocks is quite distinct, with no obvious relations in content between them; the topics mentioned in any one foregrounding block could only be related to its own intended antecedent, and not to the other. The third block of intervening material was intended to background both antecedents, by dealing with topics not directly related to either.

With the three blocks of intervening material and two alternative target sentences, each passage could be presented in any of six possible forms. As for Experiment 7, these six forms in fact represented only three conditions in a single factor design, with both target sentences able to appear in each condition. The three conditions will be termed the foregrounded case, where the intervening block and target sentence matched; the wrong foregrounded case, where the intervening block and target sentence did not match; and the backgrounded case, where either

target sentence followed the backgrounding block. Also as for Experiment 7, the assignment of passages to conditions was counterbalanced using a Latin square.

Eighteen sequences of twelve passages each were constructed. These always began with the two concealed practice passages, always in the same order, and these were followed by a random ordering of the six experimental and four distractor passages. The assignment of passages to conditions was determined by the Latin square, with the order of presentation of the passages randomised separately for each of the eighteen sequences; each subject saw one sequence. Thus, each experimental passage was presented three times in each of the six possible forms, and six times in each of the three conditions; while each subject saw all six passages, one in each form and two in each condition.

Procedure

Subjects attended a single half-hour session, during which a complete passage sequence was read. The passages were presented sentence by sentence, in exactly the manner previously described.

Results

Main Analyses

Mean target sentence reading times for the foregrounded, wrong foregrounded and backgrounded conditions were compared as a single-factor, three-level design. As each subject read two passages in each condition, the analysis was conducted on the means of these two reading times. The ANOVA summary table is given in Appendix 21.

Table 8.2

Experiment 8: Mean Pronoun Sentence Reading Times (msec)

<u>CONDITION</u>		
<u>Foregrounded</u>	<u>Wrong Foregrounded</u>	<u>Backgrounded</u>
3087	3807	4205

Table 8.2 presents the mean pronoun sentence reading times in the three experimental conditions. It can be seen that the foregrounded condition produced the fastest mean reading time, of 3087msec, while the mean reading time in the backgrounded condition was over 1100msec longer; the mean for the wrong foregrounded condition fell between these extremes, but was closer to the backgrounded condition than to the foregrounded. The analysis produced a significant result, with $F(2,34)=4.77$, $p<0.025$, and means in the three conditions were therefore compared using the Newman-Keuls procedure. For the comparison between

the foregrounded and backgrounded conditions, $Q(3)=4.31$, $p<0.05$; for the comparison between the wrong foregrounded and backgrounded conditions, $Q(2)=1.53$, not significant; and for the comparison between the foregrounded and wrong foregrounded conditions, $Q(2)=2.77$, not significant ($df=34$ in each case). It may be noted that the critical value of Q for $p<0.05$ in the last of these comparisons was 2.88, indicating that the difference between the foregrounded and wrong foregrounded conditions just failed to reach significance.

In summary, these results have strongly replicated the standard foregrounding effect found in previous experiments, and in addition have shown that the wrong foregrounded condition is no more difficult than the backgrounded condition. The wrong foregrounded condition in fact produced a mean reading time slightly faster than that for the backgrounded condition, contrary to the predictions of the working memory theory.

Subsidiary Analyses

Three additional analyses of other aspects of the experiment were also carried out, and these will now be reported. Firstly, comprehension accuracy on the two questions attached to each experimental passage was analysed, again as a single-factor, three-level design. It was not expected that experimental condition should affect comprehension performance, and this turned out to be the case, with $F<1$. Mean comprehension accuracy was 3.519, with a maximum possible of 4, indicating subjects' generally high level of performance.

Secondly, reading times for the three types of intervening material

were analysed; this procedure tested for possible differences in reading speed for intervening backgrounding material, for material foregrounding the entity mentioned first in the antecedent sentence, and for material foregrounding the entity mentioned second in the antecedent sentence. The analysis constituted a 3 x 4 design, with intervening Block and sentence Position as factors. For the main effect of Block, $F < 1$, and for the Block x Position interaction, $F(6,102)=1.15$, not significant; and for the main effect of Position, $F(3,51)=3.25$, $p < 0.05$. This position effect, which was also found in Experiment 7 but which would not reach significance here if conservative degrees of freedom were used, was again due to faster reading times for the fourth sentence position than for the first. Thus, subjects did not respond differently to the three types of intervening material, but did appear to increase in speed on later sentences.

The increased speed on later sentences was revealed again in the third additional analysis, which considered the effect of sentence serial position in the four distractor passages. The analysis was identical to that carried out in Experiment 7, in which the same distractor passages had been used. As in that experiment, the analysis produced a significant result, with $F(9,153)=11.24$, $p < 0.001$. The pattern of reading times was again similar to that found in all previous experiments; after an initially high reading time for the first sentence, reading times gradually fell until the sixth serial position, when they stabilised at between three and three and a half seconds.

Discussion

The results of this experiment seem clearly inconsistent with the working memory theory of foregrounding. Rather than producing target sentence reading times significantly longer than the standard backgrounded condition, the wrong foregrounded condition in fact produced reading times that were somewhat shorter. This conflicts with the view that wrong foregrounding should cause greater processing difficulties owing to the need to detect and cancel an erroneous immediate mapping of the pronoun before antecedent retrieval processes can be initiated. By contrast, a theory which claims that such retrieval processes are the normal method of pronoun assignment can readily account for the equivalence between the backgrounded and wrong foregrounded conditions. Retrieval processes are seen as beginning equally promptly in the two conditions, and as being subject to similar semantic and structural constraints. Since the value and distinctiveness of the working memory theory rest heavily on the assumption of immediate pronoun assignment, this experiment gives some grounds for rejecting that theory as a whole.

There are however two difficulties in accepting the conclusion that these results are incompatible with the immediate mapping strategy. The first difficulty is that the wrong foregrounded condition did not differ significantly from the standard foregrounded condition, although both the working memory theory and the retrieval theory predicted that it should. While mean reading times were clearly longer in the wrong foregrounded condition, the failure of this comparison to reach full significance implies that wrong foregrounding may have produced at least some facilitation of pronoun assignment relative to the standard

backgrounded case. Such facilitation would be difficult to explain on either theory under consideration here. It is also possible, however, that the facilitating effect of wrong foregrounding was not in fact based on pronoun assignment processes, but on other aspects of target sentence comprehension. For example, it is possible that sentences dealing with specific individuals are processed in a different way to sentences that are more general and descriptive; general content not tied to particular actions or events in a story may be processed in less depth than specific content, or may be less fully integrated into a structured representation. If readers adapt their processing strategies to the nature of the current text, it is likely that in the backgrounded case a sudden transition from general and descriptive sentences to a sentence concerned with a specific individual will cause processing difficulties as the need for different comprehension strategies is recognised. The sentence-initial pronoun will thus act as a cue signalling a change in the nature of the text, and additional processing time will be required to call up the appropriate strategies. In the wrong foregrounded case, however, comprehension of the intervening sentences presumably makes use of the same set of strategies as will be required in processing the target sentence; there will thus be no delay caused by any alteration in the manner of reading. In this way, foregrounding sentences related to one individual may produce an overall facilitation of specific references, without directly assisting pronoun assignment itself. This facilitation may have attenuated the effect of wrong foregrounding.

The second difficulty with the present results is closely related to the first. Although on the working memory theory antecedent retrieval processes should be slower to execute in the wrong foregrounded condition than in the backgrounded condition, and hence

should cause an elevation in reading times, it is evident from the above comments that total sentence reading time need not be increased if other processing consequences of wrong foregrounding actively facilitate other aspects of target sentence comprehension. One such consequence is that wrong foregrounding may avoid the difficulties involved in setting up new processing strategies to comprehend the target sentence; however, there is also a possible source of facilitation in the wrong foregrounded case which is directly concerned with pronoun assignment. In line with the strategy of immediate mapping, the sentence-initial pronoun can in the wrong foregrounded case be directly matched with an entity in working memory; this mapping will presumably be easy and rapid. In the backgrounded condition, however, an attempted direct match to working memory will fail, since no specific antecedent is currently foregrounded. It is possible that an unsuccessful attempt at direct mapping will occupy more processing time than a successful attempt. This could occur, for example, if initial failure to assign the pronoun led to a more conscious and careful examination of the contents of working memory in order to confirm that a direct match was not available, or if the awkwardness of the use of a pronoun in that context led to a delay before continuing to the remainder of the sentence. In this way, although antecedent retrieval processes will be hampered in the wrong foregrounded condition in the manner previously described, the elevation in target sentence reading times in the backgrounded condition due to failed immediate mapping may be sufficient to outweigh that effect. Target sentence reading times in the wrong foregrounded and backgrounded conditions might therefore be approximately equal, as was found here.

The danger of this alternative explanation of the present results reflects the main disadvantage of the reading time technique, which is

that it provides only a gross index of processing difficulty for each separate display, and can neither distinguish different causes of that level of difficulty nor reveal the distribution of difficulty over different sections of the display. The possibility therefore always exists that equivalent reading times in two experimental conditions may conceal very different types of processing. In the present case, the retrieval theory would seek to explain the near equivalence between the wrong foregrounded and backgrounded conditions by claiming that identical processing takes place in each; the working memory theory, however, could explain the equality by postulating that the two conditions involve different forms of processing which fortuitously produce the same overall level of difficulty. Simple comparisons of target sentence reading times do not allow these possibilities to be discriminated.

In summary, the superficial incompatibility between the present results and the suggested immediate mapping strategy of pronoun assignment depends on the assumption that equivalent reading times reflect equivalent processing. However, accepting a possibly different distribution of processing difficulty across the target sentences in different conditions would allow an explanation of these results which preserves the immediate mapping strategy. Experiment 9 attempted to discover whether or not such an explanation could be supported.

Introduction

The principal cause of the difficulty in interpreting the results of Experiment 8 was the fact that gross target sentence reading times do not directly reveal the processing which has been carried out in understanding that sentence. Reading time differences and equivalences between experimental conditions can only be accounted for in terms of predictions derived from psychological theories; in cases where two theories predict identical patterns of reading times, but do so on the basis of different types of processing, a single reading time measure is unable to distinguish between them. However, there are several techniques available which may at least reveal the distribution of processing difficulty over a sentence, rather than simply providing a gross measure for the sentence as a whole. In order to decide between the alternative explanations of the Experiment 8 result, it is the distribution of processing time which is most crucial. Under the working memory theory, antecedent retrieval processes executed at the end of the target sentence are predicted to be slower in the wrong foregrounded condition than in the backgrounded condition; but this need not lead to a corresponding difference in gross reading times if processing of the sentence-initial pronoun is sufficiently slower in the backgrounded condition.

One of the techniques which can be used to analyse the distribution of processing over a sentence is to monitor readers' eye fixations; data derived from such studies formed the basis of the reading model of Just and Carpenter (1980), which could not have been formulated or

tested so thoroughly on the basis of gross sentence reading times. The eye fixation procedure not only allows exact determination of gaze duration on each part of a sentence, but also reveals patterns of eye movements, such as regressions, which may add to the information derived from reading time alone. However, monitoring eye movements is a technically demanding procedure, and it is in fact possible to adapt the self-paced computer presentation paradigm in ways which allow more analytic investigation of sentence processing. Although the usual procedure within that paradigm is to present subjects with individual displays consisting of whole sentences, various researchers have also used presentation of units of sub-sentence size. In some cases, presentation of displays shorter than a full sentence was dictated by technical limitations on the size of the computer screen; Cirilo (1981), for example, could only present displays of up to forty characters in length, but ensured that at least the target sentences were short enough to be displayed in full. In other cases, displays smaller than a sentence have been used deliberately to investigate particular questions. For example, Aaronson and Scarborough (1976) presented their materials a single word at a time, and used the patterns of reading times over a sentence to assess the effects of syntactic structure and semantic content. The single-word technique was also used by Chang (1980), while Mitchell and Green (1978) and Green, Mitchell and Hammond (1981) presented their experimental texts in sub-sentence displays of a few words each.

The studies by Mitchell and his co-workers are especially relevant to the present questions, as they were partly concerned with whether integrative processes occur immediately a word is read or are deferred to the ends of sentences. In Mitchell and Green (1978) it was demonstrated that displays which formed the end of a sentence tended to

have longer reading times than displays which occurred earlier in a sentence; these end-of-sentence pauses were greater for sentences which had been rated as more difficult to understand, and were attributed to the fact that integrative processes were not carried out until a sentence boundary was reached. Further evidence that readers pause on the final word of a sentence has been provided by Chang (1980), as well as by Just and Carpenter (1980), who also attribute the effect to deferred information integration. This effect, and its interpretation, is clearly consistent with the suggestion of the retrieval theory that pronoun assignment processes are not normally carried out immediately the pronoun is encountered. However, Green, Mitchell and Hammond (1981) made an explicit test of whether or not certain integrative processes may be carried out as soon as a word is read, rather than being held over to the ends of sentences. In one of their studies, a text was presented in which a critical definite noun phrase, presented as a single display, either was or was not consistent with information from earlier in the same sentence; the crucial noun phrase was "the first dance", which was consistent when an earlier display referred to "a lavish formal ball" but anomalous when that display was altered to "a full council of war meeting". Clearly, in the consistent condition a referent for the target phrase could be readily calculated, while in the anomalous condition it could not. It was found that the anomalous condition not only produced longer end-of-sentence pauses than the consistent condition, as would be expected if integrative processes were deferred, but also produced longer reading times for the target noun phrase itself. This was taken to reflect the fact that readers attempted to locate a referent for the definite noun phrase as soon as it was encountered, and were more easily able to do this in the consistent condition.

The results of Green et al. (1981) are therefore in apparent agreement with the suggestion of immediate pronominal mapping which is contained in the working memory theory of foregrounding. However, there are several aspects of the Green et al. study which reduce the strength of this confirmation. Firstly, both the crucial definite noun phrase and the earlier information from which an antecedent could be computed were contained in the same sentence; the retrieval theory accepts that working memory is used to integrate sentences into wholes during comprehension, and hence the effect of a within-sentence anomaly does not distinguish between the two conflicting accounts of anaphora resolution. A more relevant test would consider the effects of an anomaly which could only be detected by comparing separate sentences. Secondly, as Green et al. recognised, the crucial noun phrase in fact formed a possible clause boundary; since it is known that clauses may also lead to integrative processes, as well as complete sentences, it is possible that readers attempted to assign the noun phrase on the basis of end-of-clause processing rather than immediate anaphor resolution. Thirdly, it is striking that Green et al. in fact used only one experimental passage, containing only one critical noun phrase; the generality of the effect which they obtained with this single instance is clearly open to question.

The feature of the Green et al. study which most weakens its applicability to the current issue, however, is the fact that it was concerned with the interpretation of definite noun phrases rather than pronouns. The possible processing differences between these two forms of anaphora have already been alluded to; in particular, the fact that definite noun phrases contain much more information than a simple pronoun makes a strategy of immediate assignment more realistic. Such a strategy in the case of definite noun phrases would not require working

memory presence of antecedents in order to succeed; although pronouns provide insufficient search cues or decision criteria to allow selection of an antecedent from the long-term text representation, definite noun phrases may be adequate guides in themselves to the intended referent. Thus, even a clear demonstration of immediate assignment of definite noun phrases would prove neither that this involved mappings to working memory nor that such a strategy was also used in pronoun assignment.

Despite these criticisms of the particular study which they carried out, the procedure used by Green et al. (1981) offers a useful and straightforward technique for investigating the current questions of interest. The working memory theory of foregrounding makes two predictions concerning the distribution of processing difficulty over the pronoun sentences used in Experiment 8. Firstly, processing at the end of the target sentence should be greater in the wrong foregrounded condition than in the backgrounded condition; this must occur, since in the wrong foregrounded case an erroneous initial assignment of the pronoun can only be detected and cancelled at the end of the sentence, while in the backgrounded case antecedent retrieval processes at the end of the sentence can execute without such a delay. End-of-sentence reading times in both conditions would be expected to be greater than in the standard foregrounded condition, where information integration on the basis of the initial pronoun assignment can proceed smoothly. The second prediction of the working memory theory is necessary to account for the fact that total sentence reading times in Experiment 8 were not in fact found to be longer in the wrong foregrounded condition than in the backgrounded condition. This prediction asserts that processing of the sentence-initial pronoun must be more difficult in the backgrounded condition than in either of the two foregrounded conditions; this would occur if successful immediate matches of a pronoun to an entity in

working memory are more rapid than failed attempts at such a match. Thus, if target sentences were presented to subjects in two sections, with the first display containing the pronoun and the second any disambiguating content, the working memory theory predicts a distinctive pattern of reading times across the three experimental conditions.

The essential difference between the predictions of the working memory theory and of the retrieval theory concerns reading times for the sentence-final display. The retrieval theory can accommodate longer reading times for the initial display in the backgrounded condition than in the foregrounded conditions, since as noted earlier these may be attributable to the change in the character of the text which is signalled by the use of a pronoun, and to the corresponding need to prepare new and more appropriate comprehension strategies; a longer reading time for the initial display in the backgrounded condition therefore need not be due to failed attempts at immediate pronoun assignment. However, the retrieval theory clearly predicts that end-of-sentence processing in the wrong foregrounded and backgrounded conditions must be identical; in both cases an antecedent must be retrieved from the text representation, and these retrieval processes will operate under similar constraints in the two conditions. Thus, the two classes of theory make distinct predictions concerning end-of-sentence reading times in the three experimental conditions, but make similar predictions concerning reading times for the initial section of the target sentences.

In order to test these predictions, this experiment made use of a presentation technique based on that employed by Mitchell and Green (1978) and Green et al (1981). The experimental passages which had been used in Experiment 8 were divided up into sub-sentence displays of a few

words each; the crucial target sentences were divided into only two displays each, as detailed in the Method section. The experiment was most concerned to discover whether reading times for the final section of the target sentences would be longer in the wrong foregrounded condition than in the backgrounded condition.

Method

Subjects

Twenty-four first-year Plymouth Polytechnic psychology students participated in this experiment in return for course credit.

Materials

As for Experiment 8, there were six experimental, four distractor and two concealed practice passages in all. These were identical to those used in the previous experiment, except for slight changes to some of the target sentences. The experimental conditions and the counterbalancing procedures were identical to Experiment 8.

The segmenting of the target sentences was carried out as follows. Each target sentence was divided into two sections; the first of these contained as much of the early part of the sentence as was considered equally compatible with either of the two crucial antecedents, while the second section contained the disambiguating content considered specific to the intended antecedent of that sentence. The initial sections were generally shorter than the final sections; the number of words in each

section varied across passages. Some slight changes were made in some cases, in order to ensure an ambiguous initial section, but the bulk of the target sentence content was identical to Experiment 8.

In order to illustrate the segmenting of the target sentences, the two targets from the first passage in Appendix 12 are given in (28). In this passage, one of the antecedent individuals was said to be gardening, while the other was said to be cleaning windows; taken as wholes, it is evident that the first sentence in (28) must refer to the man who was gardening, while the second must refer to the man who was cleaning windows. However, the initial segments of these sentences, as indicated by the slash marks, are clearly ambiguous. Thus, the sentence-initial pronoun of either sentence could be directly matched to whichever individual was currently foregrounded.

- (28) He went to fetch a // large digging fork from the shed.
He used a long // ladder to reach to the first floor.

The segmenting of all other sentences, in all passages, was carried out on a random basis. Segment length was allowed to vary from 2 to 8 words, but the randomisation was constrained in order to ensure that roughly two-thirds of the segments contained four words. Half the sentences were segmented from the beginning, and half from the end; this ensured that the longer segment lengths were able to occur approximately equally often in early and late positions. The randomisation resulted in all sentences containing 2, 3 or 4 segments; there was thus sufficient variation to make the two-segment target sentences inconspicuous. All punctuation marks were removed, except for the sentence-final full stop. Appendix 13 presents one example of the

experimental passages, with slash marks to indicate the segment boundaries. The target sentences used in this experiment are contained in Appendix 12, again with slash marks to indicate the division into initial and final segments.

Procedure

Subjects were tested individually, in a single half-hour session during which a complete passage sequence was read. The only change from previous experiments was that each display consisted of an incomplete sentence segment, rather than a full sentence. As usual, each display was centred on the middle line of the computer screen.

Subjects were warned that the displays would never form a complete sentence; that the number of displays in each sentence, and the number of words in each display, would vary; and that sentence-final full stops were the only punctuation marks present. As in previous experiments, subjects were simply told to read each display, and press the red advance button as soon as they had understood it. Successive displays were presented on the screen just as rapidly as in earlier experiments, irrespective of whether they continued a sentence or began a new one; thus, there was no noticeable pause between sentences. After the last display in each passage, the comprehension questions were presented in full exactly as in previous experiments.

Results

Main Analyses

The pronoun sentence reading time data were analysed as a 2 x 3 design. One factor consisted of the three experimental conditions, which were identical to those of Experiment 8, while the second factor consisted of the two segments into which each target sentence had been divided. These will be referred to as the initial and final segments respectively. Since each subject had read two passages in each experimental condition the analysis was conducted on the means of these reading times. The ANOVA summary table is given in Appendix 22.

Table 9.1

Experiment 9: Mean Pronoun Sentence Reading Times (msec)

	<u>EXPERIMENTAL CONDITIONS</u>			
	<u>FORE</u>	<u>WRONG FORE</u>	<u>BACK</u>	
<u>INITIAL</u> <u>SEGMENT</u>	710	863	1467	1013
<u>FINAL</u> <u>SEGMENT</u>	1401	2046	2543	2005
	1055	1454	2005	

Mean reading times for the initial and final segments and for the three experimental conditions are presented in Table 9.1. Direct comparisons between the reading times for the initial and final segments are relatively uninformative in this analysis, since initial segments

were generally much shorter than final segments; hence, although initial segments were read significantly faster than final segments, with $F(1,23)=55.20$, $p<0.001$, this need not reflect either integrative processes concerned with the pronoun nor a general end of sentence effect. The more important result for present purposes was the significant main effect of experimental condition, with $F(2,46)=6.56$, $p<0.005$. It can be seen that the mean reading time in the backgrounded condition was almost a second slower than that in the foregrounded condition, while the wrong foregrounded condition fell between these two. Differences between the three means were compared using the Newman-Keuls procedure. It was found that the wrong foregrounded condition did not differ significantly from the foregrounded condition, with $Q(2)=1.52$, nor from the backgrounded condition, with $Q(2)=2.09$. The comparison between the foregrounded and backgrounded conditions did prove significant, however, with $Q(3)=3.61$, $p<0.05$. The pattern of the main effect of experimental condition was thus very similar to that obtained in Experiment 8.

The crucial result of this experiment concerns the interaction between sentence segment and experimental condition. It can be seen from Table 9.1 that on the initial segment reading times for the foregrounded and wrong foregrounded conditions were very close together, with the reading time in the backgrounded condition being noticeably larger than either; on the final segment, however, the wrong foregrounded condition fell somewhat nearer to the backgrounded condition than to the foregrounded condition. Despite this apparent difference in pattern between the initial and final segments, the interaction did not prove significant, with $F(2,46)=1.04$. It must therefore be concluded that the differences between experimental conditions did not change significantly between segments. The most

noteworthy features of this result are that backgrounded reading times on the initial segment must be considered significantly longer than those in the foregrounded condition, while on the final segment the wrong foregrounded condition evidently did not produce a longer reading time than the backgrounded condition. This latter aspect of the results is crucial for separating the working memory and retrieval theories.

The second important analysis to be reported concerned the effect of segment position within a sentence. As was noted above, the main effect of segment on the analysis of pronoun sentence reading times may not reflect a genuine effect of position, owing to confounded differences in segment length. In order to test the effect of position, a mean reading time was calculated separately for all pre-final and all final segments read by each subject; reading times for target sentences were omitted from this calculation. Since for each subject the pre-final mean was calculated over around 270 reading times, and the final segment mean was calculated over 132 reading times, differences attributable to segment length or other factors would be expected to cancel out. It was found that mean reading time for pre-final segments was 1274msec, while mean reading time for final segments was 1379msec. This difference was in the expected direction, but did not reach full significance, with $F(1,23)=3.89$, $p<0.1$.

Subsidiary Analyses

As for Experiment 8, analyses were conducted on comprehension accuracy, reading times for the intervening sentences, and reading times for sentences in the distractor passages. For the analysis of the reading time data, a single reading time was calculated for each

sentence by summing the reading times for the several segments in that sentence. The analyses were carried out identically to those of Experiment 8.

On the analysis of comprehension accuracy, it was found that there were no significant differences between the three experimental conditions, with $F(2,46)=1.35$; mean comprehension accuracy was 3.569 out of a maximum possible of 4, and thus was very similar to accuracy levels in Experiment 8. On the analysis of reading times for the intervening sentences, no differences were found between the three types of intervening block, with $F(2,46)=2.37$; there was, however, a very slight main effect of sentence position, with $F(3,69)=2.59$, $p<0.1$. As for Experiment 8, the fourth intervening sentence was read faster than the first. On the serial position analysis of reading times for the distractor passages, a highly significant main effect of position was found, with $F(9,207)=23.38$, $p<0.001$; as for all previous experiments using the sentence by sentence display technique, reading times were faster on the later sentences than on the early sentences.

In general, therefore, the subsidiary analyses of this experiment produced patterns of effects very similar to those obtained in Experiment 8. This similarity is perhaps unsurprising, given that identical texts were used in the two experiments, yet it is interesting that the reading time results in particular appeared unaffected by the major change in presentation technique. As a check on this, an analysis was conducted to compare reading times on the distractor passages in the two experiments. This constituted a 2×10 design, with Experiments and Position as factors; data points were the means of four reading times per position for each of the eighteen subjects in Experiment 8, and means of four total sentence reading times, summed over all segments in

each sentence, for the twenty-four subjects in Experiment 9. The analysis produced the expected main effect of Position, with $F(9,360)=33.94$, $p<0.001$, but there was no main effect of Experiments, with $F<1$. The interaction was also not significant, with $F(9,360)=1.83$. It therefore appears that this large change in presentation technique did in fact produce no alterations in either reading speed or patterns of reading times over a passage.

Discussion

The results of this experiment are once again in closer agreement with the predictions of the retrieval theory of foregrounding than with those of the working memory theory. The working memory theory predicts that immediate mapping of sentence-initial pronouns to foregrounded entities should lead to added processing difficulty when that assignment is found to be incorrect; such added difficulty should produce longer reading times for the final segment of pronoun sentences. No evidence in support of this view was found here. The equivalence in final segment reading times between the wrong foregrounded and backgrounded conditions is readily explicable under the retrieval theory, which argues that pronoun assignment does not in fact involve attempted immediate matches to working memory. On this view, the required processing at the end of the target sentences would be very similar in the two conditions; in both cases, antecedent retrieval processes could execute without delay, and would operate under similar constraints.

While the pattern of reading times on the final segment was clearly incompatible with the working memory theory, the pattern of reading

times on the initial segment proved to be much as that theory would have predicted. Longer reading times in the backgrounded condition than in the foregrounded condition could be explained by arguing that in the former case the lack of any directly accessible referent led to processing delays when the immediate pronominal mapping failed. However, this reading time difference is also explicable under the retrieval theory, if it is accepted that readers adapt their comprehension strategies to the nature of the current text. In the backgrounded case, the occurrence of a pronoun indicates a transition from general and descriptive sentences to more specific and individuated topics; the need to recognise this, and call up appropriate processing strategies, may therefore lead to longer reading times than in the foregrounded condition. Hence, the reading time effect on the initial segment should not be seen as confirmation of the working memory theory. The crucial prediction which separates the two theories under consideration concerns the reading times for the final sentence segment; and these results clearly contradicted the working memory theory.

The main difficulty with these conclusions is that the interaction of sentence segment with experimental condition did not prove significant. The retrieval theory predicted equivalent final segment reading times for the wrong foregrounded and backgrounded conditions, with shorter reading times in the standard foregrounded condition; but predicted equivalent initial segment reading times in the two foregrounded conditions, with longer reading times in the backgrounded condition. Although this difference in the reading time patterns between the two segments is suggested by the means given in Table 9.1, the statistical test which would have confirmed this trend did not approach significance. Inspection of the means suggests that this was principally due to relatively faster final segment reading times in the

wrong foregrounded condition compared to the backgrounded condition. If this is in fact so, it would imply that wrong foregrounding was in some way facilitating antecedent retrieval, rather than hindering it. Such facilitation would be in even greater contradiction of the working memory theory than exact equivalence with the backgrounded condition. However, further speculation on this matter would be difficult to sustain, given the lack of significant differences between the wrong foregrounded condition and the other conditions.

These experiments have been concerned with the component of the working memory theory which asserts that anaphoric pronouns are immediately matched to foregrounded antecedents. This assertion is a central one, since to allow initially more complex assignment processes would render antecedent presence in working memory superfluous. However, the present results suggest that such immediate mappings do not in fact take place. Since in the wrong foregrounded condition the need to detect and cancel an erroneous immediate mapping should have caused greater processing difficulty than unimpeded antecedent retrieval in the backgrounded condition, pronoun sentence reading times should have been longer in the former case; no evidence for this was found, in either Experiment 8 or Experiment 9.

One possible way of explaining this result within a working memory theory would be to claim that initial direct mappings are not firmly accepted until more information is obtained to support them. Sanford and Garrod (1981), for example, suggest that preliminary mappings onto items in the working memory buffer are checked for plausibility in relation to current scenario constraints. It could therefore be argued that cancelling incorrect mappings in the wrong foregrounded condition may be relatively straightforward, since they will have been only provisionally accepted in any event, and hence reading times need not be significantly longer than in the backgrounded condition. However, it would still be the case that checking procedures would precede retrieval in the wrong foregrounded condition, while in the backgrounded case retrieval could take place without delay. It could be argued that such checking procedures themselves involved accessing long-term memory on

the basis of target sentence content, rather than simply comparing the target sentence with the current working memory contents; in this way, checking of the erroneous mapping would assist retrieval of the correct referent. However, this would contradict the assertion of Sanford and Garrod (1981), who restrict plausibility checking to currently activated scenario constraints, and would in any case be tantamount to a retrieval theory. The distinction between making assignment choices on the basis of long-term memory access, and making final confirmatory decisions on such a basis, is very weak.

Direct evidence in support of the immediate mapping theory is in fact remarkably slight. Sanford and Garrod (1981) cite the fact that semantically ambiguous pronouns may nonetheless receive generally agreed assignments as evidence that mapping precedes interpretation; this point was dealt with earlier. Sanford and Garrod also refer to certain results presented by Ehrlich (1980), which do bear on this issue. In her study of the sources of the implicit causality effect, Ehrlich discovered that sentences in which the pronoun was made unambiguous by a gender cue produced faster referent choice latencies than sentences in which there was no gender cue; in addition, altering the linking conjunction between the two clauses had no effect on referent choice when a cue was present. Ehrlich concluded that readers did not make use of general knowledge to select an antecedent when gender cues were sufficient to dictate the assignment; Sanford and Garrod (1981) agree with this conclusion. These results appear consistent with the notion that immediate mappings are attempted as soon as a pronoun is encountered.

There are two points to be made about Ehrlich's findings. Firstly, as Ehrlich herself recognised, the technique of requiring subjects to

make explicit referent choices is one which may lead to distortion of assignment strategies. While such a technique does at least ensure that assignments are being made, which the reading time technique does not, it is probable that the strategies subjects use to carry out the task will differ from those used in normal reading. One possibility is that subjects may look for the crucial pronoun as soon as each sentence is displayed, rather than reading the sentence normally; working backward from the pronoun in this way would allow an antecedent to be identified when there was a gender cue, but in the absence of such a cue the remainder of the sentence would eventually have to be read. Another possibility is that even if subjects did read the test sentences in the usual left-to-right manner, as soon as they encountered the pronoun their task goals may have overridden normal reading processes; immediate mapping when required to press a button to identify a referent is no proof of immediate mapping when simply reading for comprehension. Secondly, as noted in the literature review, in Ehrlich's experiment the gender cue was always consistent with the assignment bias signalled by either the implicit causality of the verb or the pragmatic relationships between the events described. A more revealing comparison is that presented by Caramazza et al. (1977), who found that referent choice latencies were slower when the gender cue was inconsistent with the other biases in the sentence. This implies that even in the presence of a gender cue assignment was affected by semantic constraints. The results of Hirst and Brill (1980) similarly showed effects of semantic content in cases where an assignment was determinate on syntactic grounds.

A more recent study by Ehrlich and Rayner (1983), published after the present research was completed, made an explicit attempt to discover whether pronouns are assigned as soon as they are encountered. The

rationale of this study was very similar to that of the present Experiment 9, since Ehrlich and Rayner assessed local processing difficulty on both the pronoun itself and the later parts of the sentence; this was done by recording eye fixations, however, rather than by the self-paced reading technique. Subjects were presented with texts in which the distance between a sentence-initial pronoun and its antecedent was varied. In the Near case, the antecedent was the final word of the preceding sentence; in the Intermediate case, the antecedent was an earlier noun in the preceding sentence; and in the Far case the antecedent was the first word in the last sentence but one. The crucial pronoun was always clearly unambiguous on the basis of information preceding it, notably through gender restrictions among the potential antecedents; in this way, immediacy of assignment was always possible. Fixation durations were recorded for the fixation before and the fixation during pronoun encoding, and for the two fixations following encoding. The pronoun encoding fixation was defined as that which occurred closest to the pronoun within a range of six character spaces to the left and one to the right; this was necessary, as pronouns themselves are seldom directly fixated in reading.

The principal result of interest was that while fixation durations during pronoun encoding were equivalent across the three distance conditions, fixation durations following the pronoun were substantially increased in the Far case. Ehrlich and Rayner concluded that while assignment processes may be initiated as soon as a pronoun is encountered, in the Far case they are not completed until later parts of the sentence. Such an interpretation would be fully consistent with the working memory theory of immediate mapping; only in the Far case would the antecedent not be available in working memory, and thus only in that condition would failed immediate assignment lead to continued processing

difficulty as the antecedent was retrieved while reading later parts of the sentence. Even though retrieval processes could in principle have been completed while fixating the pronoun itself, since it was clearly unambiguous at that point, the deferment of retrieval may be a usual response to failed direct mapping. However, there are a number of points to be made.

Firstly, the fact that reading times following the pronoun are longer in backgrounded conditions than in foregrounded conditions is not in itself evidence in favour of immediate mapping. Such a result was found in Experiment 9, and was easily explicable under the retrieval theory by arguing that antecedent location is simply faster in the former condition. Thus, the absence of longer post-pronoun gaze durations in the Near and Intermediate cases of Ehrlich and Rayner (1983) does not establish that no assignment processes were deferred. Secondly, if immediate mapping failed in the Far case, the working memory theory would predict longer reading times for the pronoun itself; this was argued when seeking to use the working memory theory to account for the reading time pattern on the initial segments in the present Experiment 9. No such effect was found by Ehrlich and Rayner.

This latter point is interesting, because Ehrlich and Rayner's Far case in fact constituted a wrong foregrounded case rather than a standard backgrounded case; the intervening sentence described the actions of specific individuals, different to the intended antecedent, rather than dealing with generally relevant material. In Experiment 9, a wrong foregrounded condition did not produce longer reading times for the pronoun itself, and the working memory theory could explain this by saying that rapid immediate mappings had been made to the wrong referent. However, in the Ehrlich and Rayner study such erroneous

mappings were precluded by gender constraints. Thus, their Far case should have produced longer pronoun reading times, as did the backgrounded case in Experiment 9. The fact that reading times for the pronoun itself were unaffected by distance is therefore evidence against the working memory theory. It appears that the alternative explanation offered for the initial segment effect in Experiment 9, that a sudden transition from general sentences to a sentence dealing with a specific individual led to delays due to strategy changes, may be accurate; no such delay would be present in Ehrlich and Rayner's Far case, since the intervening sentence dealt with particular individuals just as did the pronoun sentence itself.

In summary, the Ehrlich and Rayner study is consistent not only with the view that immediate mapping did not occur in the Far case, but also that it did not occur in the nearer conditions. It is interesting that this result conflicts with the findings of Ehrlich (1980), which suggested that immediate mappings based on gender did take place; this discrepancy could be explained by the unnatural strategies adopted by subjects when making explicit referent choices. The conclusions of the present experiment, that immediate mapping to foregrounded entities does not take place, is therefore not seriously contradicted by existing evidence.

CHAPTER VII

Discussion

This research has been concerned with the processes by which readers interpret anaphoric pronouns. The review of this area presented in Chapter I showed that one general issue was the manner in which pronominal antecedents appeared to vary in accessibility; certain entities in a text receive foregrounded status as a result of different influences which make them more or less available to pronominal reference. The research presented here sought both to clarify the nature of certain influences on foregrounding, and to investigate the psychological processes underlying that phenomenon. Taken together, the set of results which has been uncovered leads to a revision in theories of foregrounding, by contradicting the working memory theory described in Chapter I and agreeing with the retrieval theory presented in Chapter IV. Since the preceding description of the experimental work was accompanied by extensive discussion of each set of results, the present overall Discussion is largely intended as a summary and review of what has gone before, together with some observations concerning methodological problems and other difficulties.

Review of Experiments

The first set of experiments presented here was concerned with the effects of distance and topical status on foregrounding. The literature review showed that distance has previously been found to increase the difficulty of pronoun interpretation, but that the nature of this distance effect has been left unclear; Experiments 1 and 2 attempted to discover whether the distance effect was continuous and increasing, as suggested by Daneman and Carpenter (1980), or discontinuous in the manner suggested by Clark and Sengul (1979). Experiment 1, using an unusual presentation technique, found no effect of distance on pronoun

sentence reading times; Experiment 2 found an irregular, stepped effect on both reading times and pronoun question accuracy. It was suggested that obtained distance effects may be due to confounded variations in the topical status of antecedents, and Experiment 3 showed this to be the case. Distance was found to have no independent effect on ease of pronoun interpretation, with topical continuity in the text following an antecedent acting to preserve the antecedent's ready accessibility to pronominal reference. The lack of a distance effect in Experiment 1 may also have been due to antecedent topical status.

As noted in Chapter I, a distinction must be made between influences on assignment and processes of assignment; while Experiments 1-3 were concerned with assessing the contribution to ease of assignment of the two influences of distance and topical continuity, later experiments attempted to investigate the processes underlying those effects. These experiments concentrated on the foregrounding effect of topical continuity. In the widely accepted working memory theory of pronoun interpretation, as described in Chapter I, foregrounded items are said to be held in an immediately accessible storage system to which anaphoric pronouns are directly mapped without the use of complex sources of disambiguation (e.g. Sanford and Garrod, 1981). This theory formed the basis of Experiments 4 and 5, which used recognition memory tests to discover whether foregrounded items were in fact directly accessible to processing. Neither Experiment 4 nor Experiment 5 found any facilitation of recognition memory responses by antecedent foregrounding; this failure to obtain the predicted effect was made more convincing by the fact that Experiment 5 simultaneously replicated the foregrounding effect on pronoun sentence reading times.

The conclusion that foregrounding does not affect recognition memory was qualified, however, by the suggestion that obtaining only two observations per condition per subject may have allowed noise in the data to obscure the desired effects. This methodological difficulty was in fact general to the entire series of experiments presented here; the majority of the pronoun sentence reading time results were based on only two observations per condition per subject, although Experiment 1 did obtain three reading times. Ideally, many more reading times would have been taken in order to strengthen the power of the analyses and increase the likelihood of uncovering any genuine effects in the data. The fact that relatively small numbers of observations were obtained in each experiment raises the possibility that the pattern of results presented here may not be a true reflection of the underlying effects. However, this criticism is equally applicable to a large part of the psycholinguistic literature on reading comprehension. Daneman and Carpenter (1980), for example, made use of only two experimental passages per condition, as did the bulk of the present research; while Green, Mitchell and Hammond (1981) made use of a single experimental text. It should also be pointed out that increasing the number of passages which were presented to subjects in the present research would have increased experiment running time to an unacceptable extent; even with present technique Experiments 2-4 lasted for two hours. The procedure adopted here is therefore defensible on grounds of both practicality and conformity with published research.

A further methodological point which may be noted here concerns the use of replication to establish the generality of the obtained effects. While this procedure is recommended by Wike and Church

(1976), use of the method of analysis recommended by Clark (1973) would have had certain advantages. In particular, it would have revealed whether the obtained effects in each experiment were equally attributable to all the passages used, and hence presumably to the critical manipulation common to them all, or were only present in a certain number of passages and hence were possibly caused by unintended specific materials effects. While the basic foregrounding effect of topical continuity has been well confirmed by replication throughout the present series of experiments, this is less true of other important effects. The effect of wrong foregrounding in particular was replicated only on new subjects and under new presentation conditions; no replication on new materials has been presented. The generality of a number of results has therefore not been fully reestablished.

The absence of any foregrounding effect on recognition tests prompted a detailed reconsideration of the working memory theory. In Chapter IV it was shown that the central component of that theory was the assumption of immediate pronominal mapping; only the assumption that pronoun interpretation did not initially make use of available disambiguating content required antecedent presence in working memory to ensure successful assignment. Allowing more complex assignment processes would not only make assignment feasible without working memory presence of antecedents, but would also offer many ways of accounting for foregrounding phenomena on the basis of antecedent retrieval processes. This approach placed the emphasis on the nature of the assignment processes themselves, rather than on the appropriate prior selection of working memory contents. Assignment would be achieved more readily the more cues there were to guide the antecedent search process and the more prominently the antecedent had been coded in the long-term text representation.

It was argued in Chapter IV that the published research on foregrounding, reviewed in Chapter I, was equally compatible with both the working memory theory and the retrieval theory. That research, in common with the present Experiments 1-3, demonstrated various influences on foregrounding without making direct tests of the theory most frequently used to explain those phenomena. The present Experiments 4-5 made such a test, and produced results apparently incompatible with the working memory theory; the following experiments were specifically designed to discriminate between the predictions of the two competing explanations of foregrounding.

It was first argued that the working memory theory predicted that ease of assignment in foregrounded contexts should be unaffected by the

nature of the disambiguating content of the pronoun sentence, while the retrieval theory predicted that this content would be found to be crucial. Experiments 6 and 7 tested these two predictions. It was found that when the intended antecedent was foregrounded by means of the topical continuity manipulation, this did not produce significant facilitation of pronoun sentence reading times unless the target sentence predicate was closely related to the intervening foregrounding material. In the unrelated case, topical continuity did not succeed in maintaining antecedent availability. It was argued that this result was consistent with the retrieval theory of foregrounding, and showed that pronoun assignment made use of disambiguating sources of information even in foregrounded conditions. The result also indicated that antecedents may only be foregrounded in relation to particular pronouns in particular contexts of use, rather than being equally accessible to any subsequent pronominal reference.

Experiments 8 and 9 made a further test of the immediate mapping component of the working memory theory. It was argued that immediate mapping of pronouns to whatever entity was currently foregrounded by topical continuity should cause greater processing difficulty if the later content of the pronoun sentence showed that assignment to be incorrect; by contrast, the retrieval theory predicted equal processing difficulty in wrong foregrounded conditions and in the standard backgrounded condition. Experiment 8 showed that wrong foregrounding did not produce greater total pronoun sentence reading times than standard backgrounding, while Experiment 9 showed that wrong foregrounding also did not produce elevated reading times for the disambiguating content at the ends of the pronoun sentences. These results indicated that antecedents were selected on the basis of available disambiguating content whether or not one of the possible

alternatives was currently foregrounded by topical continuity, and that wrong foregrounding produced no more interference with these retrieval processes than neutral backgrounding. Pronouns were not immediately mapped on the basis of limited information.

The overall trend of Experiments 4-9 was evidently contrary to the predictions of the working memory theory of pronoun interpretation. Results such as those of Caramazza et al. (1977) and Hirst and Brill (1980) had already indicated that assignment processes made combined use of syntactic and pragmatic information, and the present experiments showed this to be so even in cases of antecedent foregrounding which the working memory theory argued did not require such complex processes. The retrieval theory of assignment, outlined in Chapter IV, offers a suitable alternative explanation of these results, and this is returned to shortly.

A subsidiary topic of this research was the possibility of differences between the pronoun interpretation processes of skilled and less-skilled readers. Perfetti and Lesgold (1977, 1979) suggested that poor readers may suffer from smaller effective working memory capacity in reading, while there is also evidence (e.g. Eamon, 1978-79) that poor readers are less well able to identify discourse topics. The former suggestion becomes less relevant to individual differences in sensitivity to foregrounding in the light of the preceding criticism of the working memory theory, but the latter suggestion predicts that poor readers should be less sensitive to foregrounding effects on pronoun interpretation than good readers. This prediction was tested in experiments 3, 5, and 6, all of which failed to show any evidence of skill differences in sensitivity to the effect of topical continuity. Experiments 1, 2, and 4 similarly found no differences between good and

poor readers with regard to distance effects or to effects of foregrounding on recognition memory. The lack of skill interactions in Experiments 1-6 may possibly have been due to methodological problems such as the presumably small range of abilities in the subjects used; however, the failure of the skill manipulation to produce any interesting results led to it being omitted from Experiments 7-9.

Methodological Issues

The major conclusions of this thesis are based on the evidence of pronoun sentence reading times. It has been argued that differences in reading times can be attributed to differences in the ease of pronoun assignment, and by relating these differences to the text manipulations used inferences have been drawn concerning the ways in which assignment processes operate. There are, however, a number of difficulties with the reading time technique. For example, it was necessary throughout this research to compare reading times for the different blocks of intervening sentences, in order to assess the danger of spillover effects (McKoon and Ratcliff, 1980); in addition, Experiment 9 was required because total sentence reading times do not reveal theoretically interesting aspects of the distribution of processing difficulty over a sentence. McKoon and Ratcliff (1980) make several criticisms of the reading time technique, but the fundamental difficulty is that a reading time measure does not directly reveal the processing which has been carried out. Such processing can only be inferred on the basis of theory, and these inferences can only be confirmed by repeated testing of different aspects of the theory which predicts them. This strategy was adopted here in carrying out tests of the working memory theory; while that theory could accommodate the foregrounding effect

itself, it failed to predict the other results of Experiments 4-9. However, the problem still remains that the effects obtained here have only been assumed to be due to assignment processes; as McKoon and Ratcliff (1980) point out, text manipulations may affect other aspects of processing than that which is of interest, and there is thus a danger that variations in unidentified confounding factors may be the real cause of obtained effects. There appear to be a number of factors which may have been affected by the manipulations used here, and which may have influenced sentence comprehension in ways not directly concerned with pronoun assignment.

One such possibility concerns the effects of macro-structure formation and of transitions between text constituents. It has been argued here that such processes may be partly responsible for variations in the ease of antecedent retrieval; the relevant processes were those operating soon after the antecedent sentence, which in backgrounded conditions was apparently followed by a macro-boundary. However, it is possible that in backgrounded conditions target sentence reading times may have been elevated by macro-processes occurring during comprehension of that sentence itself. In backgrounded conditions, the target sentence indicates closure of the current topic and reversion to an old topic; as such, these sentences may be accompanied by macro-processes operating to reduce the content of the intervening block of backgrounding sentences. They may also form category transitions, which are known to produce longer reading times (Mandler and Goodman, 1982). Neither of these influences would operate in foregrounded conditions. Thus, at least part of the longer reading times in backgrounded conditions may have been caused by macro-processes which are not involved in assigning the target pronoun.

A second possible source of longer pronoun sentence reading times in certain conditions than in others may be the greater difficulty of integrative processes which occur after the pronoun has been assigned. It was noted earlier that research on pronoun interpretation has centred largely on the assignment process rather than on the combining of information which assignment makes possible; locating an antecedent is seen as the first step required for information integration, but effects on pronoun interpretation are normally attributed to the former process rather than to the latter. Post-assignment integration processes may, however, offer one explanation for the relatedness effect of Experiments 6-7. Assuming, contrary to the present conclusions, that assignment was achieved equally easily in the related and unrelated foregrounded conditions, in the latter case comprehension might be hampered by the need to integrate information concerning one aspect of the antecedent with a representation which was largely coded in terms of some other attribute. However integration proceeds, this inconsistency would presumably cause difficulties which need not be part of the assignment process itself.

Other possible sources of reading time differences unconnected with the processes of assignment include subjects' recognition of the awkwardness or incomprehensibility of the target sentences (McKoon and Ratcliff, 1980), and the possible priming effect of semantic closeness between context and target which was mentioned in connection with Experiments 6-7. A problem which increases the difficulty of assessing the likelihood of these alternative sources of the obtained reading time effects is that no guarantee can be given that the crucial pronouns were in fact being assigned. Subjects were instructed to read each sentence and press the advance button as soon as they understood it; however, it is possible that full understanding was not always attempted.

Experiment 6 showed that the foregrounding effect was reduced in the latter half of a sequence of experimental texts, and this result is compatible with the suggestion that under some circumstances subjects accepted relatively shallow comprehension as sufficient. Hence, it is possible that in some conditions subjects may not have attempted to assign the crucial pronoun, and that had they done so the pattern of results would have been changed. This is relevant particularly to the effects of wrong foregrounding found in Experiments 8-9; if in the wrong foregrounded condition subjects curtailed their attempts at comprehension of the target sentence as soon as the initial, directly-mapped assignment proved incorrect, then reading times would not in fact have been increased compared to the standard backgrounded condition. Such an increase was predicted under the working memory theory only on the assumption that readers would go on to attempt retrieval of the correct antecedent once their initial assignment was seen to be erroneous; if they did not do so, the reading time equivalence between backgrounded and wrong foregrounded conditions would be compatible with the working memory theory.

The two techniques which have been used to ensure that readers do assign the crucial pronouns, or to ascertain whether such assignments have been made, are the pronoun question technique and the explicit referent choice technique. However, the former has already been extensively criticised here, while the latter may clearly distort subjects' reading strategies to an unacceptable extent. In general, no single research technique should be considered ideal for purposes of investigating pronoun interpretation; important results should preferably be tested on more than one methodology. The advantage of doing so was shown in the present research, when comparisons between the reading time technique and the pronoun question technique found both

similarities and differences between the results they gave. Perhaps more importantly, the present research has obtained results based on reading time data which argue against the direct mapping theory of pronoun assignment; similar conclusions may also be drawn from the results of Hirst and Brill (1980) and Caramazza et al. (1977), who employed the referent choice technique.

The replication of important results on different methodologies may be one way of assessing the effects of confounded factors such as those mentioned above. For the present, however, these dangers can only be noted here. While such alternative sources of reading time effects may possibly contaminate the present results, such dangers are endemic to psychological research and can only be left for fuller investigation at some future time.

Pronoun Interpretation Processes

The view of pronoun interpretation being argued for here is in clear contrast to views such as those presented by Sanford and Garrod (1981). In common with other proponents of the working memory theory of immediate pronominal mapping, these authors argue that initial attempts at pronoun assignment do not make use of constraints derived from pragmatic or semantic information; these restricted assignment processes succeed if the antecedent is foregrounded, and only in backgrounded conditions are complex search and decision processes called into play. The retrieval theory, however, sees assignment as routinely making use of combined sources of information which are jointly assessed in the process of locating the required antecedent. The assignment process is sensitive to structural relationships between pronoun and

antecedent as well as to pragmatic relationships, and variations in the quantity and quality of these different types of information can be used to account for foregrounding phenomena.

An interesting consequence of the retrieval theory concerns the issue of whether foregrounding is a continuum or an all-or-none dichotomy. In the literature review, it was noted that certain results suggested that foregrounding might vary continuously; in particular, results presented by Sanford and Garrod (1981) which showed effects of distance or of scene changes to be overlaid on effects of global topicality seemed consistent with this view. It was pointed out in Chapter IV, however, that results such as those of the present Experiment 3 indicated foregrounding to be an all-or-none distinction; this is also a view consistent with results such as those of Clark and Sengul (1979) and Lesgold et al. (1979). It was also pointed out in Chapter IV that the explanation for a continuum of foregrounding offered by Sanford and Garrod (1981), who allege that items in working memory receive different allocations of storage space, undermined the principal motivation for the working memory theory as a whole. The retrieval theory, however, offers a simple explanation for continuous variations in levels of foregrounding which is based on essential features of the theory. Since the theory attributes antecedent foregrounding to the ease of operation of standard retrieval processes, it can be clearly seen that such processes should show a continuum of speed and efficiency. This was implied in the general discussion of Experiments 6-7, which argued that the backgrounding effect might be weakened if more information was provided to act as an antecedent retrieval cue. The results of Experiments 6-9 in fact support the view that foregrounding is a continuum, since they showed unrelated and wrong foregrounded conditions to fall between the standard foregrounded and

backgrounded conditions.

The retrieval theory has been developed here primarily to account for antecedent foregrounding in cases where pronoun and antecedent are not closely adjacent in the surface text. Some types of intervening material appear to push antecedents into the background, while others act to preserve antecedent accessibility to pronominal reference. None of Experiments 7-9 included a condition with no intervening sentences, while Experiment 6 did so principally in order to replicate the interaction between distance and foregrounding found in Experiment 3. However, the relationship between assignment processes in the non-adjacent foregrounded conditions, and in conditions where the pronoun sentence immediately follows the antecedent sentence, is an interesting problem. It seems plausible that in cases of closely adjacent pronouns and antecedents the antecedent will in fact be present in working memory when the pronoun is encountered; one component of the retrieval theory accepts that information may be carried over from the preceding sentence to assist in integration of the next, while in cases of within-sentence anaphora the simultaneous working memory presence of pronoun and antecedent would be still more likely. If this is so, it might be expected that working memory presence of antecedents in adjacent conditions should facilitate pronoun assignment; Experiments 3 and 6, however, showed that pronouns in adjacent conditions were no easier to process than in non-adjacent foregrounded conditions. Superficially, this may appear to suggest that, since antecedents are present in working memory in adjacent conditions, they must also be present there in non-adjacent conditions in order to account for the unchanged ease of assignment.

It is clear, however, that even in adjacent conditions pronoun assignment makes use of combined sources of different types of information. This was shown by the results of Hirst and Brill (1980), who found that semantic plausibility affected ease of pronoun assignment to syntactically constrained antecedents in the preceding sentence, and by the results of Caramazza et al. (1977), who found that inconsistent verb biases hampered within-sentence assignments which were constrained by gender cues. The need for use of available disambiguating content in adjacent cases is also indicated by the observation, made earlier, that working memory may contain representations of several entities in addition to the desired antecedent. Thus, even if the intended antecedent is present in working memory, as it may be assumed to be in closely adjacent conditions, pronoun assignment evidently does not take place by means of direct mapping based solely on the content of the pronoun itself. Since the present research shows that assignment in non-adjacent foregrounded conditions also does not take place by direct mapping, pronoun assignment processes can be seen as operating in a relatively uniform manner in all foregrounded and backgrounded conditions.

Nevertheless, since the retrieval theory attributes the foregrounding effect to differences in the ease of operation of the standard assignment processes, it might be argued that working memory presence of antecedents in adjacent conditions should similarly speed the execution of those processes. This suggestion, however, rests on the assumption of strictly sequential processing which characterised the working memory theory. Working memory presence might be expected to facilitate pronoun interpretation if the assignment processes made an initial search of working memory and only subsequently went on to consider the long-term text representation; there is, however, no

empirical basis for this assumption, and it runs counter to the frequent assertion that different reading processes may operate simultaneously. Thus, lack of reading time differences between adjacent and non-adjacent conditions could be explained by arguing that searches of working memory and long-term memory operated in parallel. It was noted earlier that mappings to working memory could not be intuitively guaranteed to be more rapid than retrieval searches, and that the working memory theory was therefore forced to assume serial and additive processing stages in order to confidently predict slower assignment times in backgrounded conditions; and it appears still more plausible that retrieval searches of long-term memory in foregrounded conditions should be as rapid as searches of working memory in adjacent conditions.

It may, however, be inaccurate to regard searches of working memory and searches of long-term memory as distinct processes, whether executed serially or in parallel. Since searches of the long-term text representation are presumably carried out by the central processing component of the working memory system, and since the retrieval theory assumes that these searches are guided in part by the contents of working memory, it appears that working memory searches and long-term memory searches may be different aspects of a single processing operation. Encountering a pronoun may lead to readers executing uniform assignment processes in all contexts; these processes will involve drawing on different sources of disambiguation and accessing subsections of long-term memory. The identification of an antecedent which was already present in working memory would occur as part of these general assignment processes, rather than as a result of a specific working memory search. Working memory presence need not facilitate antecedent location if the processes of assignment are not specifically directed to look there, but only discover such presence while executing standard

retrieval processes.

The above observations suggest that working memory presence of antecedents in adjacent conditions need not facilitate assignment compared to non-adjacent foregrounded conditions. However, it should be pointed out that the evidence from the present research may be considered somewhat equivocal as to whether or not there is a distance effect in foregrounded conditions. The two investigations of this question were made in Experiments 3 and 6. In Experiment 3, the crucial result was the interaction between distance and foregrounding, which was interpreted as showing the presence of a distance effect in the backgrounded case but the absence of a distance effect in the foregrounded case. However, when full degrees of freedom were used this interaction reached only the $p < 0.05$ significance level, and with conservative degrees of freedom the interaction missed significance by a large margin. In Experiment 6, the relevant interaction did not reach significance even when using full degrees of freedom, and the main evidence for the desired effect came from the one-way analysis of variance which made direct comparisons between distance 0 and the other distance conditions. The one-way analysis was also used to confirm the effect in Experiment 3. While such an analysis avoids the problems of random variations between the dummy foregrounding conditions at distance 0, the interaction on the original analysis would be the most desirable proof that distance has no effect if topical continuity is preserved.

On balance, given the results of the one-way analyses and considering the fact that conservative degrees of freedom offer a very stringent criterion for significance, the present research does indicate that there is no strong or reliable distance effect in foregrounded conditions. However, the possible existence of a weak and elusive

effect gives some support to the view that non-adjacent foregrounded antecedents are not located in working memory. The principal direct evidence for this claim comes from the present Experiments 4 and 5, which failed to obtain a foregrounding effect on recognition memory tests; since it was noted there that antecedents occurring in sentence-initial position in the antecedent sentence need not be in working memory immediately after that sentence was read, this provides another possible partial explanation of the lack of distance effect in foregrounded conditions.

The present research therefore supports a general view of pronoun interpretation which stresses the active, antecedent-seeking nature of the assignment process. Rather than being relatively automatically assigned to whatever entity is currently most available in working memory, pronouns are interpreted by means of complex search and decision processes which draw on several sources of information. These information sources include aspects of sentence and text structure as well as semantic cues derived from the immediately preceding sentence and the pronoun sentence itself. The combination of the various information sources allows a full specification of the required antecedent to be constructed, and thus assists not only the antecedent search process but also decision processes which confirm that the antecedent has been found. Pronoun assignment will be made more difficult in cases where the different pieces of information are inconsistent. Ease of antecedent location is affected by the appropriacy of the semantic cues used to guide the search, the position of the antecedent in the text hierarchical structure, and the extent of the macro-structure reduction of the antecedent information. Topical continuity in the material following mention of an entity prevents macro-structure reduction, provides valuable semantic retrieval cues,

and maintains the current text in a close structural relation to that entity, all of which acts to preserve its accessibility to pronominal reference. In foregrounded cases, the search process is thus readily able to identify the required antecedent in the long-term text representation. In backgrounded cases, however, the relative absence of guiding semantic cues and the inconsistency between certain cues and others, together with macro-structure reduction of the antecedent information and greater structural separation between pronoun and antecedent, all act to hamper the assignment process.

This description of pronoun assignment processes does not purport to be either precise or complete. As was noted in the Introduction, the concept of working memory which has been applied to reading has been left vague and unspecified; because of this, theories of pronoun interpretation based on such a concept would be difficult to test in any detail. The present research, however, has succeeded in discriminating between two general classes of pronoun assignment theory, largely by showing that predictions of the working memory theory which were independent of the specific form of that store were incorrect. The inadequacy of direct mapping to working memory as an explanation of foregrounding has been demonstrated, while an explanation based on antecedent retrieval processes has been shown to offer a plausible alternative. Several suggestions were made in Chapter IV concerning how these processes may operate; these suggestions must however be put to empirical test in a more fully specified form before any firm conclusions on their exact nature can be drawn.

REFERENCES

- AARONSON, D., and SCARBOROUGH, H. S. (1976) - Performance theories for sentence coding: Some Quantitative Evidence. Journal of Experimental Psychology. Human Perception and Performance. 2, 56-70
- AKMAJIAN, A., and JACKENDOFF, R. (1970) - Coreferentiality and stress. Linguistic Inquiry, 1, 124-126
- ANDERSON, J. R., and BOWER, G. H. (1973) - Human Associative Memory. New York. Wiley.
- ATKINSON, R.C., and SHIFFRIN, R.M. (1963) - Human Memory: A Proposed System and Its Control Processes. In K.W. Spence and J.T. Spence (eds) - The Psychology of Learning and Motivation, Vol. 2. New York: Academic Press.
- BADDELEY, A.D. (1979) - Working memory and reading. In P.A. Kolars, M.E. Wrolstad and H. Bouma (eds) - Processing of Visible Language, Vol. 1. New York: Plenum
- BADDELEY, A.D. (1981) - The concept of working memory: A view of its current state and probable future development. Cognition, 10, 17-23.
- BADDELEY, A., ELDRIDGE, M., and LEWIS, V. (1981) - The role of subvocalisation in reading. Quarterly Journal of Experimental Psychology, 33A, 439-454.
- BADDELEY, A.D., and HITCH, G. (1974) - Working Memory. In G.A. Bower (ed) - The Psychology of Learning and Motivation, Vol. 8. New York: Academic Press
- BADDELEY, A.D., and LIEBERMAN, K. (1980) - Spatial working memory. In R.S. Nickerson (ed) - Attention and Performance VIII. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- DE BEAUGRANDE, R. (1980) - Text, Discourse and Process. London: Longman.
- DE BEAUGRANDE, R., and DRESSLER, W.U. (1981) - Introduction to Text Linguistics. New York: Longman.
- BLACK, J.B., and WILENSKY, R. (1979) - An evaluation of story grammars. Cognitive Science, 3, 213-229
- BRANSFORD, J.D., BARCLAY, J.R., and FRANKS, J.J. (1972) - Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 3, 193-209.
- CAPLAN, D. (1972) - Clause boundaries and recognition latencies for words in sentences. Pneftian and Psychophysics, 12, 73-76.
- CARAMAZZA, A., GROBER, E., GARVEY, C., and YATES, J. (1977) - Comprehension of Anaphoric Pronouns. Journal of Verbal Learning and Verbal Behaviour, 16, 601-609.

- CARPENTER, P.A., and JUST, M.A. (1977a) - Reading comprehension as eyes see it. In M.A. Just and P.A. Carpenter (eds) - Cognitive Processes in Comprehension. Hillsdale, N.J.: Lawrence Elbraum Associates.
- CARPENTER, P.A., and JUST, M.A. (1977b) - Integrative processes in comprehension. In D. La Berge and S.J. Samuels (eds) - Basic Processes in Reading: Perception and Comprehension. Hillsdale, N.J.: Lawrence Elbraum Associates.
- CARR, T.H. (1981) - Building theories of reading ability: On the relation between individual differences in cognitive skills and reading comprehension. Cognition, 9, 73-114.
- CHAFE, W.L. (1972) - Discourse structure and human knowledge. In J.B. Carroll and R.O. Freedle (eds) - Language Comprehension and the Acquisition of Knowledge. Washington, D.C.: V.H. Winston and Sons.
- CHAFE, W.L. (1976) - Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C.N. Li (eds) - Subject and Topic. New York: Academic Press
- CHANG, F.R. (1980) - Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. Memory and Cognition, 8, 58-64.
- CHOMSKY, N. (1981) - Knowledge of Language: Its elements and Origins. Philosophical Transactions of the Royal Society of London B, 295, 353-374.
- CIRILO, R.K. (1981) - Referential coherence and text structure in story comprehension. Journal of Verbal Learning and Verbal Behaviour, 20, 358-367.
- CIRILO, R.K., and FOSS, D.J. (1980) - Text structure and reading time for sentences. Journal of Verbal Learning and Verbal Behaviour, 19, 96-109.
- CLARK, H.H. (1973) - The Language-as-fixed-effect Fallacy: A critique of language statistics in psychological research. Journal of Verbal Learning and Verbal Behaviour, 12, 335-359.
- CLARK, H.H., and HAVILAND, S.E. (1974) - Psychological processes as linguistic explanation. In Cohen, D. (ed) - Explaining Linguistic Phenomena. Washington, D.C.: Hemisphere Publishing.
- CLARK, H.H., and HAVILAND, S.E. (1977) - Comprehension and the given-new contract. In R.O. Freedle (ed.) - Discourse Production and Comprehension. Norwood, N.J.: Ablex
- CLARK, H.H., and SENGUL, C.J. (1979) - In search of referents for nouns and pronouns. Memory and Cognition, 7, 35-41.
- DANEMAN, M., and CARPENTER, P.A. (1980) - Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behaviour, 19, 450-466.

- VAN DIJK, T.A. (1977a) - Semantic macro-structures and knowledge frames in discourse comprehension. In M.A. Just and P.A. Carpenter (eds.) - Cognitive Processes in Comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- VAN DIJK, T.A. (1977b) - Text and Context. London and New York: Longman
- VAN DIJK, T.A., and KINTSCH, W. (1977) - Cognitive Psychology and Discourse: Recalling and summarizing stories. In W.U. Dressler (ed.) - Current Trends in Text Linguistics. Berlin: Walter De Gruyter.
- DUTKA, J.T. (1980) - Anaphoric relations, comprehension and readability. In P.A. Kolers, M.E. Wrolstad, and H. Bouma (eds.) - Processing of Visible Language, Vol. 2. New York: Plenum
- EAMON, D.B. (1978-79) - Selection and Recall of Topical Information in Prose by Better and Poorer Readers. Reading Research Quarterly, 14, 244-257.
- EDWARDS, A.L. (1972) - Experimental Design in Psychological Research. New York: Holt, Rinehart and Winston.
- EHRlich, K. (1980) - Comprehension of pronouns. Quarterly Journal of Experimental Psychology, 32, 247-255.
- EHRlich, K., and RAYNER, K. (1983) - Pronoun assignment and semantic integration during reading: Eye movements and immediacy of Learning. Journal of Verbal Processing and Verbal Behaviour, 22, 75-87.
- FLETCHER, C.R. (1981) - Short-term memory processes in text comprehension. Journal of Verbal Learning and Verbal Behaviour, 20, 564-574.
- FODOR, J.A., BEVER, T.G., and GARRETT, M.F. (1974) - The Psychology of Language. New York: McGraw-Hill.
- FREDERIKSEN, J.R. (1981) - Sources of Process Interactions in reading. In A.M. Lesgold and C.A. Perfetti (eds.) - Interactive Processes in Reading. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- GARROD, S., and SANFORD, A. (1977) - Interpreting anaphoric relations: The integration of semantic information while reading. Journal of Verbal Learning and Verbal Behaviour, 16, 77-90.
- GARROD, S.C., and SANFORD, A.J. (1981) - Bridging inferences and the extended domain of reference. In J. Long and A.D. Baddeley (eds.) - Attention and Performance IX. Hillsdale N.J.: Lawrence Erlbaum Associates.
- GARVEY, C., and CARAMAZZA, A. (1974) - Implicit causality in verbs. Linguistic Inquiry, 5, 459-464.

- GARVEY, C., CARAMAZZA, A., and YATES, J. (1976) - Factors influencing assignment of pronoun antecedents. Cognition, 3, 227-243.
- GOLDMAN, S.R., HOGABOAM, T.W., BELL, L.C., and PERFETTI, C.A. (1980) - Short-term retention of discourse during reading. Journal of Educational Psychology, 72, 647-655.
- GRAESSER, A.C., HOFFMAN, N.L., and CLARK, L.F. (1980) - Structural components of reading time. Journal of Verbal Learning and Verbal Behaviour, 19, 135-151
- GREEN, D.W., MITCHELL, D.C., and HAMMOND, E.J., (1981) - The scheduling of text integration processes in reading. Quarterly Journal of Experimental Psychology, 33A, 455-464.
- GROBER, E.H., BEARDSLEY, W., and CARAMAZZA, A. (1978) - Parallel function strategy in pronoun assignment. Cognition, 6, 117-133.
- GROSZ, B.J. (1981) - Focussing and description in natural language dialogues. In A.K. Joshi, B.L. Webber, and I.A. Seg (eds) - Elements of Discourse Understanding. Cambridge: Cambridge University Press.
- HABERLANDT, K. (1980) - Story grammar and reading time of story constituents. Poetics, 9, 99-118.
- HABERLANDT, K., and BINGHAM, G. (1978) - Verbs contribute to the coherence of brief narratives: Reading related and unrelated sentence triple. Journal of Verbal Learning and Verbal Behaviour, 17, 419-425.
- HALLIDAY, M.A.K. (1970) - Language structure and language function. In J. Lyons (ed.) - New Horizons in Linguistics. Harmondsworth: Penguin Books.
- HAVILAND, S.E., and CLARK, H.H. (1974) - What's new? Acquiring new information as a process in comprehensions. Journal of Verbal Learning and Verbal Behaviour, 13, 515-521.
- HIRST, W.H., and BRILL, G.A. (1980) - Contextual aspects of pronoun assignment. Journal of Verbal Learning and Verbal Behaviour, 19, 168-175.
- HOBBS, J.R. (1982) - Towards an understanding of coherence in discourse. In W.G. Lehnert and M.H. Ringle (eds.) - Strategies for Natural Language Processing. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- JARVELLA, R.J. (1971) - Syntactic processing of connected speech. Journal of Verbal Learning and Verbal Behaviour, 10, 409-416.
- JOHNSON-LAIRD, P.N. (1981) - Comprehension as the construction of mental models. Philosophical Transactions of the Royal Society of London B, 295, 353-374.
- JUST, M.A., and CARPENTER, P.A. (1980) - A theory of reading: From eye fixations to comprehension. Psychological Review, 87, 329-354.

- KIERAS, D.E. (1981) - Component processes in the comprehension of simple prose. Journal of Verbal Learning and Verbal Behaviour, 20, 1-23.
- KINTSCH, W. (1974) - The Representation of Meaning in Memory. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- KINTSCH, W. (1977) - On comprehending stories. In M.A. Just and P.A. Carpenter (eds.) - Cognitive Processes in Comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- KINTSCH, W., and VAN DIJK, T.A. (1978) - Toward a model of text comprehension and production. Psychological Review, 85, 363-394.
- KINTSCH, W., KOZMINSKY, E., STREBY, W.J., MCKOON, G., and KEENAN, J.M. (1975) - Comprehension and recall of text as a function of content variables. Journal of Verbal Learning and Verbal Behaviour, 14, 196-214.
- KINTSCH, W., and VIPOND, D. (1980) - Reading comprehension and readability in educational practice and psychological theory. In L.G. Nilsson (ed.) - Perspectives in Memory Research. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- KINTSCH, W., and YARBROUGH, J.C. (1982) - Role of rhetorical structure in text comprehension. Journal of Educational Psychology, 74, 828-834.
- KOLERS, P.A., WROLSTAD, M.E., and BOUMA, H. (1980) - Processing of Visible Language, Vol. 2. New York, Plenum.
- LESGOLD, A.M., ROTH, S.F., and CURTIS, M.E. (1979) - Foregrounding effects in discourse comprehension. Journal of Verbal Learning and Verbal Behaviour, 18, 291-308.
- LEVY, B. (1978) - Speech analysis during sentence processing: Reading and listening. Visible Language, 12, 81-102.
- LEWIS, D. (1979) - Scorekeeping in a language game. Journal of Philosophical Logic, 8, 339-359.
- LYONS, J. (1968) - Introduction to Theoretical Linguistics. Cambridge: Cambridge University Press.
- MANDLER, J.M. (1978) - A code in the node: The use of a story schema in retrieval. Discourse Processes, 1, 14-35.
- MANDLER, J.M., and GOODMAN, M.S. (1982) - On the psychological validity of story structure. Journal of Verbal Learning and Verbal Behaviour, 21, 507-523.
- MANDLER, J.M., and JONSON, N.S. (1977) - Rememberance of things parsed: Story structure and recall. Cognitive Psychology, 9, 111-151.

- MARATSOS, M.P. (1973) - The effects of stress on the understanding of pronominal coreference in children. Journal of Psycholinguistic Research, 2, 1-8.
- McKON, G., and RATCLIFF, R. (1980) - The comprehension processes and memory structures involved in anaphoric reference. Journal of Verbal Learning and Verbal Behaviour, 19, 668-682.
- MEYER, B.J.F., BRANDT, D.M., and BLUTH, G.J. (1980) - Use of top-level structure in text: Key for reading comprehension of ninth-grade students. Reading Research Quarterly, 16, 72-103.
- MILLER, J.R., and KINTSCH, W. (1980) - Readability and recall of short prose passages: A theoretical analysis. Journal of Experimental Psychology: Human Learning and Memory, 6, 335-354.
- MITCHELL, D.C. (1982) - The Process of Reading. Chichester: Wiley.
- MITCHELL, D.C., and GREEN, D.W. (1978) - The effects of context and content on immediate processing in reading. Quarterly Journal of Experimental Psychology, 30, 609-636.
- NELSON, H.E., and WARRINGTON, E.K. (1980) - An investigation of memory functions in dyslexic children. British Journal of Psychology, 71, 487-503.
- OAKHILL, J. (1982) - Constructive processes in skilled and less skilled comprehenders' memory for sentences. British Journal of Psychology, 73, 13-20.
- PERFETTI, C.A., FINGER, E., and HOGABOAM, T. (1978) - Sources of vocalisation latency differences between skilled and less skilled young readers. Journal of Educational Psychology, 70, 730-739.
- PERFETTI, C.A., and GOLDMAN, S.R. (1976) - Discourse memory and reading comprehension skill. Journal of Verbal Learning and Verbal Behaviour, 14, 33-42.
- PERFETTI, C.A., and HOGABOAM, T. (1975) - Relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology, 67, 461-469.
- PERFETTI, C.A., and LESGOLD, A.M. (1977) - Discourse comprehension and sources of individual differences. In M.A. Just and P.A. Carpenter (eds.) - Cognitive Processes in Comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- PERFETTI, C.A., and LESGOLD, A.M. (1979) - Coding and comprehension in skilled reading and implications for reading instruction. In L.B. Resnick and P.A. Weaver (eds.) - Theory and Practice of Early Reading, Vol. 1. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- RUMELHART, D.E. (1975) - Notes on a schema for stories. In D.G.

- Bobrow and A. Collins (eds.) - Representation and Understanding: Studies in Cognitive Science. New York: Academic Press.
- RUMELHART, D.E. (1977) - Toward an interactive model of reading. In S. Dornic (ed.) - Attention and Performance VI. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- SACHS, J.S. (1967) - Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 2, 437-442.
- SANFORD, A.J., and GARROD, S.C. (1980) - Memory and attention in text comprehension: The problem of reference. In R.S. Nickerson (ed.) - Attention and Performance VIII. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- SANFORD, A.J., and GARROD, S.C. (1981) - Understanding Written Language. Chichester, Wiley.
- SCHANK, R., and ABELSON, R.P. (1977) - Scripts, Plans, Goals and Understanding. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- SHELDON, A. (1974) - The role of parallel function in the acquisition of relative clauses in English. Journal of Verbal Learning and Verbal Behaviour, 13, 272-281.
- SIDNER, C. (1979) - Disambiguating references. In P.H. Winston and R.H. Brown (eds.) - Artificial Intelligence: An MIT Perspective, Vol 1. Cambridge, Mass.: The MIT Press.
- STANOVICJ, K.E. (1981) - Attentional and automatic context effects in reading. In A.M. Lesgold and C.A. Perfetti (eds.) - Interactive Processes in Reading. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- THORNDYKE, P.W. (1977) - Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, 9, 77-110.
- VIPOND, D. (1980) - Micro- and Macroprocesses in text comprehension. Journal of Verbal Learning and Verbal Behaviour, 19, 276-296.
- WEBBER, B.L. (1980) - Syntax beyond the sentence: Anaphora. In R.J. Spiro, B.C. Bruce and W.F. Brewer (eds.) - Theoretical Issues in Reading Comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- WIKE, E.L., and CHURCH, J.D. (1976) - Comments on Clark's "The Language-as-fixed-effect fallacy". Journal of Verbal Learning and Verbal Behaviour, 15, 249-255.
- YEKOVICH, F.R., WALKER, C.H., and BLACKMAN, H.S. (1979) - The role of presupposed and focal information in integrating sentences. Journal of Verbal Learning and Verbal Behaviour, 18, 535-548.

APPENDIX 1

EXPERIMENTAL AND PRACTICE PASSAGES -

EXPERIMENT 1

A. EXPERIMENTAL PASSAGES

- (1) The restaurant was very popular.
 Many tables were almost full.
 More people were still arriving.
 A waiter was alone in the rush.
 He/A man protested to the manager.
 Some people had been left unserved for a long time.

 Were customers getting good service?
- (2) It was the final day of the race meeting.
 The racecourse was very busy.
 The stands were filling rapidly.
 A jockey fell at the second fence.
 He/A man cried out in shock.
 One of the following horses also fell.

 Did a large number of people see the accident?
- (3) The golf links were quite bush.
 The course was very difficult.
 Bunkers lined the narrow fairways.
 A golfer hunter through some long grass.
 He/A man had lost a ball.
 The search had gone on for some time.

 Was it easy for players to set low scores?

(4) The football match was well attended.
 The floodlights were shining brightly.
 The fixture would shortly begin.
 A policeman moved slowly through the crowd.
 He/A man had been knocked down earlier.
 There were bound to be more incidents.

 Had there been trouble before the match?

(5) It was a busy market day.
 Various goods were on display.
 Crowds bustled thickly all around.
 A shopper scanned the people with care.
 She/A woman had just been robbed.
 The pickpocket might still be nearby.

 Had the thief been caught?

(6) Fire engines soon appeared at the scene of the blaze.
 Flames flickered through every window.
 Columns of smoke billowed upwards.
 A fireman pulled one of the hoses.
 He/A man was ready to turn on the water.
 The fire would be easily controlled.

 Did the firemen take long to reach the fire?

- (7) The chemist's shop stocked the usual things.
 Sticking plasters stood near shampoos.
 Toiletries partly filled the shelves.
 An assistant got out some headache pills.
 She/A woman had complained of a bad migraine.
 The tablets would soon cure it.
- Were the hair shampoos kept close to the plasters?
- (8) The reference library was on the first floor.
 Desks were of brown mahogany.
 Fitted shelves reached the ceiling.
 A librarian went to fetch a volume.
 She/A woman wanted to check a reference.
 There were few other people there.
- Were the tables made of metal?
- (9) The garage was on a main road.
 The forecourt was always busy.
 A good workshop handled repairs.
 A mechanic moved an old tyre aside.
 He/A man wanted drive a car in.
 The petrol tank was leaking.
- Did the garage have a good location?

(10) The traffic jam stretched for miles.
 Cars waited bumper to bumper.
 Impatient drivers blew their horns.
 A policeman stood very near the kerb.
 He/A man was brushing by a car.
 The slow speed meant no damage was done.

 Were drivers being held up for long periods?

(11) The antique shop was cramped.
 Furniture crowded the small space.
 Shelves carried less bulky items.
 An assistant moved a small oak table.
 He/A man wanted to get to a cabinet.
 There was very little room to move,

 Did the shop have a great deal of stock?

(12) The film had already begun.
 Sweet wrappers were being rustled.
 Smoke from cigarettes rose gently.
 An usherette flicked on a small flashlight.
 She/A woman was looking for an empty seat.
 There were few places left.

 Was the film poorly attended?

(13) The hospital ward was very noisy.
Breakfast was just being finished
Cups and saucers were rattling.
A nurse called out loudly for help.
She/A woman had slipped and fallen badly.
A doctor had to be called.

Was the fall serious?

(14) The doctor's waiting room was crowded.
Health notices covered the walls.
Appointments were running quite late.
A receptionist fetched a small plastic cup.
She/A woman needed a drink of water.
There was very little talking.

Were patients being seen on time?

(15) The hotel had an impressive entrance hall.
The floor was richly carpeted.
Well upholstered chairs stood nearby.
A clerk rang a small silver bell.
He/A man wanted a suitcase carried up.
Porters were always on hand.

Was the hotel an expensive one?

(16) The factory had good security precautions.
 Alsation dogs were on patrol.
 High fencing surrounded the works.
 A guard pushed a gate half open.
 He/A man was about to drive a van in.
 The gates were locked at night.

 Would the factory be easily broken into?

(17) The road had been dug up.
 Trenches stretched across the pavement.
 Earth heaps prevented free movement.
 A labourer yelled out a loud warning.
 He/A man had almost been hit by a bulldozer.
 Such accidents were common.

 Could pedestrians walk along the footpath as usual?

(18) The charity parade was going well.
 Hot sunshine attracted a crowd.
 There were many colourful floats.
 A bystander got out a few coppers.
 He/A man called to a passing collector.
 Many people were making donations.

 Were a large number of people watching the parade?

B. PRACTICE PASSAGES

(1) The street was almost empty.

It was a dreary afternoon.

The rain was falling steadily.

A traffic warden came out of a sidestreet.

He hurried to a parked car.

It was standing on double yellow lines.

Was the vehicle illegally parked?

(2) Holidaymakers filled the railway station.

The platforms were very crowded.

A porter snatched up a small suitcase.

Late passengers hurried for trains.

A man had almost left it behind.

There was a great deal of confusion.

Had all the travellers reached the station in good time?

APPENDIX 2

EXPERIMENTAL PASSAGES - EXPERIMENT 2

A. DISTANCE O

(1) When I reached the coffee shop that Saturday most of the others were there.

I had to explain about washing my father's car for some extra pocket money.

We all started talking about how we managed on the money we were given.

Being at school had lots of drawbacks, and poverty was one of them.

Tom told us he had given up his paper round because of the low pay.

Mick and Johnny declared that the newsagents were only exploiting schoolkids.

As the talk went on, my attention drifted, and I stared out of the window.

It struck me as pointless to complain about things we could not alter.

My reverie was broken when more of our friends came noisily down the stairs.

Larry was there, but so were the girls, including Linda and Maureen.

They had been playing pinball in the upstairs room for almost an hour.

Karen was laughing about the way she and Sue had interfered with Jeff's game.

I moved up the bench, making room for Barbara, who had brought some coffees.

When I looked out of the window again, I saw Clare coming up the street.

She waved, but carried on walking past the cafe.

- 1) Who walked past the cafe?
- 2) How were newsagents exploiting schoolchildren?
- 3) Where were the pinball machines?

- (2) They spent most of that Saturday afternoon working in their large garden.

The house was new, and there was a lot to be done.

Though it was hard work, the young married couple enjoyed it.

After a few hours, they came inside and Pat started to prepare a meal.

She was peeling some potatoes when the knife slipped unexpectedly.

There was little pain, but it still gashed her wrist quite badly.

Her husband came and helped her to wash the cut under the tap.

He fetched the first-aid kit, but the bleeding refused to stop.

In the end, they became worried enough to telephone for an ambulance.

At the hospital, the ambulanceman helped her down out of the vehicle.

A porter held open the door, and inside a nurse asked what was wrong.

Before much time had passed a doctor came and looked at the injury.

He could see that the cut would need several stitches.

- 1) Who saw that the cut would need stitching?
- 2) Where were the couple working that afternoon?
- 3) What was Pat doing when she cut herself?

- (3) Variety shows are one of the few thriving forms of theatre entertainment.

They have something for everyone, which ensures their popularity.

Seat prices are usually low, and the night out can be cheap.

Mary and Jim often liked to spend an evening out in this way.

The most recent show they had seen had really been very good.

The first performer was a comedian, who soon livened up the audience.

After that came a conjuror, and then there was a ventriloquist.

A singer was top of the bill, appearing last.

Her act was very well received by the crowd.

As the theatre emptied, people talked about the show they had seen.

The manager stood in the foyer, listening to their comments as they passed.

In this way he got some idea of which acts had been a success.

Any artist making a good impression was given a return engagement.

- 1) Whose act was well received by the crowd?
- 2) Was a night out at a variety show expensive?
- 3) What did the manager do as the audience was leaving?

- (4) In the week before the election, campaigning became hectic and intense.

Each night, volunteer canvassers were given areas of the town to cover.

With leaflets and arguments, they tried to win people's votes.

Neil stayed in the committee room, giving instructions.

The room was in a state of constant upheaval, reflecting the busy activity.

In the car park outside, there was the sound of a car door slamming.

In a few moments, Ken, Reg, and Martin came in, with Anthony close behind.

Gordon was already sitting in front of the fire getting warm.

He moved over to make room for the new arrivals.

It was a cold, damp night, and no-one was enjoying the campaign.

Once they had finished their area, people came back for a cup of tea.

Suddenly the wind howled more loudly, as a strong gust rattled the windows.

The outside door flew open and banged loudly against the wall.

- 1) Who moved over to make room for the new arrivals?
- 2) Why did the outside door fly open?
- 3) Were people paid to go canvassing?

B. DISTANCE 2

(5) It was the Friday before the summer break, and we were all in good spirits.

The factory was quiet, only a couple of people still working for their bonus.

No-one else did much work that afternoon, and the foremen accepted it.

I joined a poker game with Don, Norman, and Harry, but soon gave up.

It seemed pointless to lose the money for my holidays before they even began.

When we finally clocked out, there was a big rush to get through the gates.

I stopped on the far pavement, talking to Jim, Bob, and Sid.

When Les came over, we all arranged to meet that night for a drink.

Pete started to cross the road, not seeing a lorry turning the corner.

The driver sounded his horn, but it was too late to do any good.

The brakes were full on, but the road surface there was very oily.

The truck skidded and hit him hard on the back.

1) Who was hit by the truck?

2) Why did the truck skid?

3) Why were some people still working that afternoon?

- (6) The quietest section of the carpet mill was the finishing department.

Most of the work in that area was done by hand, mainly by women.

In the distance, the roar of the heavy looms could barely be heard.

The women sat around tables mending small flaws in rugs.

It was delicate, precise work, but after a while it became very boring.

Sometimes the radio would be switched on to help pass the time.

The best distraction, of course, was being able to talk to the others,

Hilda, Christine, and Bessie always sat at the same bench.

Margaret, another of the group of friends, would usually be there too.

Jane was the youngest girl, and always had to make the tea.

As the morning wore on, the women would look forward to the first break.

Sometimes they came to work without breakfast, and wanted a hot drink.

She would put the kettle on five minutes before the break.

- 1) Who would put the kettle on?
- 2) Which department did the women work in?
- 3) What was the best distraction from the boring work?

- (7) Every Easter Saturday, the Pace Egg drama was performed in the town centre.

A group of amateur players put on the show for charity, on a street corner.

It was an old tradition, and generally drew a big audience.

The dragon was a favourite character, with his green cardboard head.

A man in drag played the fair maiden, getting most of the laughs.

St. George looked on as the evil knight pretended to attack the crowd.

Audience participation was a great feature of the performance.

People had to cheer and boo at the right places in the action.

Children shrieked in mock terror as he swung a huge mace.

After the show, the players would pass round collecting boxes.

Anyone trying to sneak off was soon spotted, and booed at loudly.

This meant that almost everyone donated something, for the local charities.

Often the group gave special shows at hospitals or children's homes.

- 1) Who swung a huge mace?
- 2) Which character got most of the laughs?
- 3) When was the collection made?

- (8) The town hall housed not only the council offices, but also the county court.

An usher directed a man to the courtroom, where a case was under way.

The trial had paused, as the admissibility of some evidence was challenged.

This could clear the accused, but the prosecution wanted it disallowed.

The prosecutor stated his arguments precisely, in a superior tone of voice.

The defending barrister was infuriated, and replied angrily and to the point.

While this went on, the defendant sat quietly, apparently unconcerned.

A reporter from the local newspaper took notes of what was being said.

After a time, the judge asked the jury to leave the courtroom.

This was obviously necessary at the stage of the argument.

The jury could not be allowed to hear the evidence being debated.

His request was quickly obeyed by the jury members.

In the absence of the jury, the wrangle became slightly less heated.

Outside the court, the jurors waited to be called in again.

They speculated about what this vital piece of evidence could be.

It was an interesting case, and it would probably be hard to reach a verdict.

- 1) Whose request was quickly obeyed?

- 2) What infuriated the defending barrister?
- 3) Where was the county court located?

C. DISTANCE 4

- (9) The campus was covered in notices announcing the big debate. It was in the Great Hall that night, and had attracted a lot of attention.

The hall was unusually full, well before the start of the discussion.

The atmosphere was stifling because of the large numbers.

A steward had to go around opening all the windows.

Eileen and Hazel arrived late, because they had been playing squash.

Eileen was an excellent player, and the match had been completely one-sided.

They went and sat with Sharon and Debbie, who had saved seats for them.

Their friend Tessa was standing at the lectern, making a speech.

The debate had reached fever pitch, and there was a lot of heckling.

Some opponents of the motion began to clap their hands slowly.

There were boos and catcalls from the other side of the hall. To make things worse, there was a fault with the loudspeaker system.

Her small voice could not be heard above the din.

- 1) Whose voice could not be heard above the din?

2) Who had won the squash match?

3) Where was the debate held?

(10) It had turned out to be a marvellous day for a walk.

The Somerset countryside was bathed in sunshine, with only a slight breeze.

Three friends made their way slowly across the fields.

They had been walking for hours, and were eager to get home.

Though it had been fun, they were tired and a little footsore.

They came to a narrow river, which barred their path.

John suggested that they should swim across, carrying their clothes.

Mike agreed that it would be quicker than walking to a bridge.

Only Tom, who was a poor swimmer, was afraid.

It is true that beginners are wiser to stay in swimming pools.

Any river could turn out to have an unexpectedly strong current.

There might be hidden branches which could catch a leg.

There would be no trained lifeguards nearby to help.

In the end he decided the danger was too great.

1) Who decided the danger was too great?

2) In what part of England was the passage set?

3) Had the friends enjoyed their walk?

(11) Janet had invited some school friends round to listen to records.

Her parents had bought her a new stereo, and she wanted them to see it.

They spent the whole evening experimenting with the various controls.

It was a thoroughly enjoyable few hours, but had to end some time.

As the bus-stop was just outside, they could wait in the warmth of the hall.

Sally, Sheila, and Jenny had their coats buttoned up, ready to dash out.

Tina and Lesley were in the living room, taking a last look at the stereo.

Paula went down to the gate to see if the bus was coming.

It was early in December, and the night air was very chilly. Overhead, the stars were shining brightly in the clear wintry sky.

The exhausts of passing cars left ghostly trails of white vapour.

There was hardly a breath of wind, and it would be a bitterly cold night.

Her feet kept slipping on the frost that was already underfoot. When the bus arrived, they all went upstairs and sat near the back.

The conductor got angry with them because he had no change.

In the end, though, they got their tickets and started talking.

Janet was a bit of a show-off, but on the whole they decided they liked her.

- 1) Whose feet kept slipping on the frost?
- 2) Did the girls have the correct fare ready?
- 3) Who had bought Janet the stereo?

(12) Stein's Menagerie toured Central Europe at about the turn of the century.

It had sideshows and stalls, but the main attraction was its zoo of predators.

Stein, the owner, knew his business, and kept only the fiercest carnivores.

Old Hans went ahead of the show, advertising the wild beasts to be seen there.

People in the small villages listened wide-eyed to his fearsome stories.

They would eagerly pay to see for themselves, and were never disappointed.

With its huge mane, the lion seemed strong and proud, truly a king of beasts.

The fine coats of the leopard and cheetah seemed to conceal their ferocity.

The South American puma, all in black, had an air of mystery. People's blood was really chilled with terror when they saw the Bengal tiger.

Now, people in those regions may be more liable to flights of fancy than most.

They do have a traditional belief in the occult, in vampires and ghouls.

Nevertheless, there was something uncanny in this powerful sense of dread.

In the village inns, small groups would discuss the matter.

The villagers felt sure that it was an evil animal.
 People from the show began to sense the growing hostility.
 Stein became displeased when the crowds visiting the
 show began to thin.
 Eventually, he was forced to turn back to more civilized
 regions.
 In places where the population was less superstitious,
 business would improve.

- 1) Which animal did the villagers feel sure was evil?
- 2) When did Stein become displeased?
- 3) Which supernatural beings did the villagers traditionally
 believe in?

D. DISTANCE 6

- (13) There was a slight mist that Sunday morning, but the sun
 was getting through.
 It would be a grand day for the planned fishing trip.
 They had all set their alarms the night before, and had
 an early start.
 Adam drove his car, loaded with their gear, right up to
 the river bank.
 Gus and Nigel walked a little upstream, while Colin went
 the other way.
 Ray was the only one without a fold-up seat, and had to
 sit on the ground.
 As the sun shone more brightly, wisps of steam rose up from
 the water.

There were small splashes as fish rose occasionally to the surface.

The river was in a beautiful area, with trees and fields all around.

It was a perfect place for a day out, and soon it would be quite hot.

There was a small pub a mile down the road which served good lunches.

Most important of all, the fishing was good and they soon began to bite.

Only the wet dew on the grass spoiled his day.

- 1) Whose day was spoiled by the wet grass?
- 2) What woke the men up that morning?
- 3) What day of the week was it?

- (14) There was a great deal of activity on the University campus that day.

Degrees were to be conferred on the new graduates in the main hall.

On the stage, red leather chairs had been provided for the dignitaries.

People began to collect outside a little while before the ceremony began.

Tony had got a Second, but Dave had done very well, getting a good First.

Martin and Frank were making fun of him, but were secretly envious.

Jerry came running up the street, his gown flapping wildly.

The academic gowns did look strange on the young people.

Because they were usually hired, a good fit was rarely to be found.

In any case. they generally gave an impression of fancy dress.

Some people thought that such dressing up should be made optional.

However, all attempts to end the tradition had failed.

There was still something compelling in the ancient ritual.

He reached the hall feeling too out of breath to speak.

1) Who was too out of breath to speak?

2) What special chairs had been provided?

3) Did the academic gowns often fit well?

(15) Jack and Ruth had three daughters and two sons, all married with children.

A Christmas, the whole clan came to visit, and Ruth cooked a special meal.

She was an excellent cook, and her Christmas lunches were always marvellous.

By tradition, they all sat around the big table in order of their ages.

At the foot, and at a small side table, were the young grandchildren.

Graham sat opposite his wife Louise, about half-way along the table.

Then came Anna and her husband Clive, and Jane with Stuart.

Ellen and James came next, while Alan, the eldest son, was

near the head.

Carol, his wife, sat opposite him, and would help by serving out the soup.

The chatter before the meal actually began was light and happy. There were always new stories to be told about the growing children.

Someone might praise the Christmas decorations, and everyone would agree.

It made the family feel closer and more secure to sit together in this way.

Private jokes, which no outsider could understand, marked their familiarity.

They would not begin to eat until everyone had been served. Because of this she filled each bowl quickly and efficiently. Later, Jack would carve the turkey, taking great pride in his expertise.

The turkey was big enough to let everyone have a good portion.

There were tureens of vegetables, and fat pork sausages. After the meal, no-one felt like moving very much for at least an hour.

- 1) Who filled the bowls quickly and efficiently?
- 2) How many children did Jack and Ruth have?
- 3) What feature of the chatter marked the family closeness?

(16) Mrs. Viner did not want to go shopping that day, but she had no choice.

It was the end of the month, and all the bills had to be paid.

In the morning it rained, but when it cleared in the afternoon she set off.

First she called at the television showroom, to pay the rental.

The Gas Board and Electricity Board were next, on the same street.

Then she went to get some stamps, and finally did some shopping.

There was only a small queue at the butcher's, and none at the fruit shop.

As it was nearly closing time she then hurried towards the bank.

She walked quickly along the pavement, clutching two large bags.

In her haste, she bumped into someone, and her shopping went flying.

It took several minutes to pick it all up, some of it covered in dirt.

At the place where she had to cross the road, she stood waiting on the kerb.

There was a large puddle nearby, and a car drove straight through it.

Her coat was covered in muddy brown marks and blotches.

Her day was complete when she found it was already closed.

She turned around and set off homeward, feeling very depressed.

The bags were heavy, and she had a long walk up a steep hill.

When she got home, the children were back from school, and

made her some tea.

She sat down for a few minutes, but soon they would want some food.

- 1) Which place did Mrs. Viner find was already closed.
- 2) Why was Mrs. Viner's coat covered in brown marks?
- 3) When did the rain clear?

E. DISTANCE 8

- (17) In June, I was invited to spend an evening at Sir Henry's country house.

I accepted at once, as he was a good host and always served fine wine.

The dinner was excellent, and afterwards we moved into the music room.

It was old fashioned, but the family always had live music in the evenings.

Lady Blythe sat on a sofa with her husband, Freddy.

Muriel took the empty place between them, as the seats filled up.

There were comfortable armchairs for Desmond, Patrick, and Simon.

Steven was the first to perform for us, playing a short piece on the violin.

The Maria played the cello, and Kenneth a brisk flute tune.

After this, George sat down at the piano and began to play.

I had always loved this instrument, and hoped to hear something special.

As the first bars of a Chopin waltz floated on the air,
I leaned forward.

This was one of my favourite pieces, and I knew it by heart.
Through the windows, the sounds of the country mingled with
the music.

The evening air and the still bright sky made a perfect
setting.

All of Chopin's tenderness and melancholy seemed contained
in every note.

I felt myself caught up, giving all my attention to the sounds.
Only very seldom do I find myself moved to such an extent.
There was no doubt that his playing was superb.

- 1) Whose playing was superb?
- 2) Who sat next to Lady Blythe?
- 3) What was the time of year?

- (18) The streets of the small Spanish town thronged with people.
Normally a sleepy village, at fiesta time it bustled with life.
The bullfights were the big attraction, and visitors packed
the hotel.
- It was a famous spectacle, and people even came from abroad
to watch.
- At a pavement cafe, a group of English tourists gathered.
James and his wife Laura were bullfight fans, and came
every year.
- Albert was also a regular visitor to the festivities, enjoying
them a lot.
- Charles was there for the first time, and was unsure what
to expect.

Crowds pressed against the thin wooden railing of the cafe. The streets were filled with a constant hum of excited voices.

Some local children had already begun to let off firecrackers. Ole people winced unhappily at the loud unexpected bangs. The real celebrations of the fiesta would not begin until the next day.

In a traditional test of their courage, young men would race ahead.

The fights themselves would begin late in the afternoon. He hoped the killing would not be too unpleasant to watch.

- 1) Who hoped the killing would not be too unpleasant to watch?
- 2) Where were the group of English tourists sitting?
- 3) How would young men test their courage?

(19) On the day of the local league snooker final, Frank Worth was taken ill.

Harry, Don, Ian, and the other members of his team were frantic.

Sam Brayshaw, the league chairman, ruled they could pick a substitute.

Only Jeff, a total novice, was available at such short notice. On the face of it, the team's prospects were not very good. The opposing side were very strong, and naturally felt self-confident.

Snooker is not an easy game, and it takes time to learn it well.

As with all sports, real skill only comes after long practice. The spectators realised the match would now be less entertaining. A few people grumbled as they stood around the playing area. Both teams were nevertheless determined to play a good game. Neither side would admit that the result was a foregone conclusion.

He might after all turn out to have a lucky game.

The first frame of the match began as the clock struck eight.

By the time of the first interval, the scores were level.

The conversation around the bar was lively and excited.

Everyone was getting caught up in the finely balanced contest.

- 1) Who might turn out to have a lucky game?
- 2) Where did the conversation take place during the interval?
- 3) At what time did the match begin?

- (20) In the large conference room, the directors gathered for a board meeting.

The treasurer talked to the managing director as the coffee was served.

The company secretary had begun taking his without sugar

People were always making fun of his round and portly figure.

There was a tense atmosphere as they waited for the chairman to get there.

Everyone knew that this meeting would decide the firm's foreseeable future.

Because of the general economic recession, profits were falling rapidly.

There was a proposal to close one of the plants, to concentrate resources.

If this happened, even jobs at top management level would be at risk.

This meeting would be a stormy one, since the plan was fiercely opposed.

A group of executives had consistently argued for a new product development.

This would offer long-term security, rather than just short-term relief.

The battle-lines were firmly drawn, and no-one could predict the outcome.

When he finally arrived, people quickly took their places.

The different factions at on opposite sides of the table.

After a few opening remarks, the discussion centred on the proposed closure.

Tempers grew heated as each side showed statistics to support their claims.

At times, the arguments almost flared into personal insults.

- 1) Who finally arrived?
- 2) Why had the company secretary stopped taking sugar?
- 3) Why were the company's prpfit's falling?

APPENDIX 3
PRACTICE AND DISTRACTOR PASSAGES -
EXPERIMENTS 2-5

Where appropriate, each passage is followed by the recognition test sentence presented in Experiment 2.

Each passage is also followed by the single recognition test word presented in Experiment 4.

A. EXPLICIT PRACTICE PASSAGES

(1) The old people's home stood in several acres of woods and gardens.

Surrounded by a high stone wall, it was a quiet and restful place.

Mrs. Shipman had lived there for almost two years, and was quite content.

She had only one child, a son, who was a senior executive in a London bank.

She quite understood that he was too busy to look after an aging woman.

He had at least paid a lot of money for her to live here in comfort.

The home was well-run, and staffed by friendly, helpful nurses.

There was always plenty to do, with walks in the garden or books to read.

The other residents were very nice, and Mrs. Shipman had made new friends.

It was really preferable to spending all day alone in her son's house.

Before she had moved to the home, she hardly met or talked to anyone.

Now, when her son came visiting, she could assure him she was very happy.

She only wished he could get away from his work more than once a month.

1) How long had Mrs. Shipman lived at the home?

2) Did the home charge high fees?

3) How often did Mrs. Shipman's son visit her?

home

(2) On Sunday, Karen took her dog, Rex, for a walk in the park, Rex was a large, boisterous Alsation, and needed plenty of exercise. Walking through the streets to the park, Karen kept the dog on a tight lead.

She had tried to train him to walk to heel, but he had a mind all of his own,

Given half the chance, he would dash into the road, chasing motorbikes. In the park, it was safe to let him run free for a while.

At first, Karen just walked along, letting Rex run around exploring. Then she got out a rubber ball, and threw it for him to chase. He would run after the ball all right, but often refused to bring it back.

He would stand over it, looking at Karen expectantly, until she ran up to him.

She often wondered if the dog thought he was taking his owner for a walk.

When the time came to go home, she had to put the lead on quickly. If he saw it coming, he would run away, and have to be caught.

- 1) What breed of dog was Rex?
- 2) Why did Karen keep him on a lead in the streets?
- 3) What did Karen throw for him to chase?

It was safe in the park to let him run free for a while

stick

B. CONCEALED PRACTICE PASSAGES

(1) Up on the terminal roof, the view of the runways was marvellous.

Telescopes were provided for the watchers, to look at the aircraft.

Mrs. Manders and her son Peter had been there for an hour, watching it all.

He was fascinated by anything to do with aeroplanes.

The only problem was the noise from the jets as they taxied just below.

At last, Mrs. Maunders said they had to go back down into the building.

They went into the observation lounge, where Mr. Maunders was waiting.

They had driven out to the airport just so Peter could see the aircraft.

In the lounge, the noise was greatly reduced by good double glazing.

Mr. Maunders went to the bar for some drinks and a packet of crisps.

Peter ate the crisps mechanically, watching a plane taxi out for take-off.

His father said this would be his last chance, as they had to go soon.

When the plane vanished in the distance, they all stood up to leave.

1) Who was fascinated by anything to do with aeroplanes?

2) Why was it quieter in the observation lounge?

3) Where were the telescopes?

Simon

(2) The Cadstone and District Loyalty Band was a local institution.

Every family in the small community had someone connected with it. For such a small town, they still produced large numbers of good musicians.

Jerry and Steve learned the trumpet at school, and joined the band later.

Like many of the players, they now worked at the local textile mill.

The firms owning the mill had once offered to sponsor the band.

A stormy meeting was held in the village hall when the proposal was made. Though giving financial security, the band's name would have been changed.

In the end, the proposal was voted down, and the band kept its old title.

Now it depended on concert entrance fees for financial survival.

However, the strong local support meant its existence was in no danger.

A sign of prosperity was the new set of uniforms that had been bought.

These were for the national brass band contest to be held later in the year.

Musical skill was the main factor, but presentation was also noticed.

1) Where did Jerry and Steve work?

2) Why was the offer to sponsor the band voted down?

3) Why had new uniforms been bought?

trumpet

(3) Mrs. Warren prepared for her annual visit to her daughter with great care.

Despite going there each year, she never got used to the trip.

First she packed a large suitcase, with clothes and things for her stay.

Then she filled the smaller bag, with all the items for the train. Magazines, sandwiches, flask, paper tissues, travel sickness pills. She was nearly seventy, so perhaps her ritual precautions were excusable.

A taxi took her to the station, and the driver helped with her bags. Then there would be a porter to carry her big case onto the platform. She always tried to find a seat facing the engine.

It was a long trip, and she had to change trains on two occasions. She would grow nervous and fidgety long before the train pulled in. She would rush across the platforms, even if she had half an hour to spare.

Her worst nightmare was of being stranded by a missed connection. At other times, she would worry about derailments and collisions. She was also anxious about pickpockets, clutching her handbag tightly. When her daughter met her at the station, she always looked worn out. Yet she had made the same journey for years, and always seemed to survive. To be strictly truthful, she probably enjoyed worrying so much.

- 1) How old was Mrs. Warren?
- 2) What was her worst nightmare?
- 3) How long did she usually spend at her daughter's?

thief

(4) It still seemed like the middle of the night when John's alarm clock woke him.

Everything was pitch black, but he jumped out of bed almost at once. Milkmen have to get used to getting up at the crack of dawn.

His breakfast things were laid out ready on the table, to save time. He could never face work without a hot meal inside him.

Driving to the depot, he noted with satisfaction that the sky was clear. There was a slight frost, but the weather he really hated was rain.

At the depot, he loaded his van as quickly as he could, and then set off.

The streets were dark and empty, but lights were coming on in the houses.

John had to collect the money that day, as well as deliver the milk. The round always took longer, as no-one ever had the right money ready.

He called every week at the same time, but they always seemed surprised.

Standing waiting on the doorstep in the cold was a real nuisance.

- 1) How did John get to the depot?
- 2) What season of the year was it?
- 3) What kind of weather did John really hate?

Lights were coming on in the houses, but the streets were drab and empty.

bacon

(5) In the autumn, Graham liked to take early morning walks along the cliffs.

During the summer, there were too many tourists for it to be much fun.

He would start out just as it was getting light, and make for North Head.

This was a large headland just outside the town.

The path wound its way around the edge of a sheer drop to the sea below.

It was dangerous if there was a strong wind, as you could be blown over.

From the headland, Graham would follow the path down onto the beach.

He would usually look for driftwood, or anything left by the tide.

One day, he found the beach littered with seabirds covered in oil.

There were patches of oil on the sand, and more still floating in.

He ran at once to call the police and report his discovery.

A tanker had obviously dumped some oil ballast illegally, just offshore.

This had happened once before, and hundreds of birds had died.

- 1) Why was it not much fun to walk along the cliffs in summer?
- 2) What did Graham usually look for on the beach?
- 3) Was the oil from an accidental spillage?

headland

(6) Unwin's was the only large department store in the centre of the town. It had a seven-floor building to itself, in the main shopping area. It was owned by a local firm, and had a long tradition of good service. By offering high quality at low prices, it defeated all competition. None of the national department store chains dared open a branch there. Like all large stores, Unwin's lost a lot of money through shoplifting. John Langton, the manager, considered this to be his worst problem. Close-circuit television had been installed, on his recommendation. To the store detective, Mike Carson, this was an inadequate precaution. He trained his staff to mingle with the crowds, and keep their eyes open. More thieves were caught that way than by fancy new gadgets. He secretly considered the television system to be a waste of money. Once they were caught, thieves were taken to the manager's office. The store's policy was to always prosecute people found stealing.

- 1) Who considered the television system to be a waste of money?
- 2) Were thieves ever let off with a warning?
- 3) Was the store owned by a national chain?

screen

(7) Lorna and Brian had booked a weekend holiday in London for themselves. Their children had been sent to stay with Ivy, Lorna's elder sister. On Friday they set off, driving to the railway station with their luggage. Because it would be only a short stay, they had just one small case each. When they arrived at King's Cross, they took a taxi to the hotel. Their room was just as the brochure had promised, with its own bathroom. They were eager not to waste any time, so they soon made off into town. That evening they contented themselves with having a meal at a restaurant. On Saturday, though, they set off in earnest to see the sights. They had never been to London before, and enjoyed finding their way about. They made a tour of the shops, and bought a few things. After lunch, they went to see Buckingham Palace and Westminster Abbey. In the evening they had been booked to go and see a show at a theatre. They arrived back at the hotel in the small hours, tired but happy. On Sunday they simply walked around, looking at the usual tourist spots. Back at home, they told their friends what a wonderful trip it had been.

- 1) Who looked after the children that weekend?
- 2) What did they do on their first evening?
- 3) How did they get to the hotel from the station?

(8) The Winston Players were a popular local amateur dramatic group.

There were always good crowds when they performed their shows.

Generally they only attempted light comedies, which people always liked.

The plays were staged in the civic hall, which was now a hive of activity.

Final preparations were being made for their latest production.

Various volunteer helpers were putting the finishing touches to the scenery.

Though they were enthusiastic, they were hopelessly disorganised.

Andy, the producer, had to keep them under constant supervision.

Terry and Charlie held up an imitation fireplace, waiting for orders.

No-one seemed to be quite sure where on the stage it was supposed to go.

There was just as much confusion in the small rooms backstage.

The costumes were all too big, and were having to be altered.

The people doing the refreshments rushed around with cups and teapots.

With each new crisis, the success of the show seemed more and more unlikely.

It was like this every time, but somehow things always turned out all right.

- 1) Who was the producer of the show?
- 2) Where were the plays staged?
- 3) Why were the costumes being altered?

coffee

(9) At seven o'clock, Glenda left her house and walked round to the next street.

She stopped at a house half-way along, and rang the doorbell.

The door opened almost at once, and Susan came out with her coat already on.

Glenda should have called half an hour earlier, as they had arranged.

They were going into town to see a new film at the cinema.

As they walked, they talked casually about work and people they knew there.

When they reached the cinema, they had an argument about where to sit.

Glenda wanted to go in the balcony, but Susan preferred to stay downstairs.

The ticket seller began to grow impatient as the girls held up the queue.

She began to tap on the counter with her ballpoint pen.

In the end, Glenda got her way and they bought tickets for the balcony.

Upstairs, they found seats and settled down, as the lights went dim.

They had only just arrived in time, as the film was about to start.

Susan was uncomfortable upstairs, and was almost sorry she had come.

She was even more irritated by Glenda's habit of talking during the film.

She never stopped commenting on the acting or the locations.

But still the two were good friends, despite their disagreements.

- 1) Who tapped on the counter with a pen?
- 2) Where did the girls sit?
- 3) Why were they nearly late for the film?

balcony

(10) The allotments at the end of Vicar Street were very much in demand.

People applying for one had to go at the bottom of a long waiting list.

The patches of land were well tended, often with neat little

sheds nearby.

The locals gave them a lot of attention, growing flowers and

vegetables.

It was a disaster when one morning they were found to have been

vandalised.

Mr. Johnson, who lived nearby, discovered the damage, and called

the police.

Then he telephoned all the people he knew who had allotments there.

Ernie Townsend and Jack Wilkes came down to see for themselves.

The place looked as though it had been hit by a bomb.

Plants had been torn up, greenhouses smashed, and sheds broken into.

Some tools had been stolen, but the general destruction was more serious.

The police took statements and looked around, but they were not

hopeful.

With no witnesses, the chances of catching the culprits were very slim.

It was astonishing that no-one had heard the noise during the night.

1) Who telephoned the police?

2) What had been stolen?

3) Where were the allotments?

Plants had been torn up, sheds broken into, and greenhouses smashed.

tomato

(11) Jerry went to the launderette very early that morning as he always did.

He found doing his washing a nuisance, and liked to get it over with quickly.

Going early meant that he had the place to himself, and left the day free.

It was a short walk, and he hurried along, carrying a bag full of laundry.

The main road was already busy, with people on their way to work, As usual, though, the launderette was empty when he got there.

After putting his things in the machine, Jerry looked out of the window.

He spent a few minutes watching the street, and then went for a newspaper.

The newsagent's was just next door, and he was back in a moment.

Soapy water was pouring out of the washing machine, flooding the floor.

He ran to open the lid, as the water began to turn black and oily.

He looked in, and saw that his clothes had been ruined by the oily mess.

Jerry swore to himself, but there was nothing to be done.

He decided to wait for the attendant, to make his complaint in person.

His money certainly ought to be refunded, at the very least.

- 1) Where was the newsagent's?
- 2) Why was the main road busy?
- 3) Why did Jerry go to the launderette so early?

coin

(12) Sue had worked at the public library near her home for just on four years.

She had gone there straight from school, after taking her A levels. Before then, she had worked there part-time, on Saturdays and some evenings.

When she first began, she had imagined a quiet, bookish sort of life. In fact, she was quickly forced to see that this was pure fantasy. A librarian's work may once have been like that, but things had changed.

The library was a centre not just for book lending, but for many activities.

Children's story evenings, links with adult education, information services.

Most of Sue's time was spent dealing with people, answering their questions.

Surrounded by books as she was, she hardly had time to read any of them.

Yet she enjoyed the job, mainly because she liked the people she worked with.

The senior librarians were very helpful, and the atmosphere was relaxed.

The younger staff were all great friends, and got on very well.

On a practical level, career prospects were good and the job was secure.

Eventually, she hoped to become head librarian in a library of her own. She could already think of some innovations she might like to bring in.

- 1) What was the main reason why Sue enjoyed the job?
- 2) How long had she worked at the library?

3) What was most of her time spent doing?

shell

(13) Michael Palmer was a middle-rank executive in the firm of Garner and Sons.

He had worked up to that position over the years, and felt he did a good job.

When they were taken over by a larger French firm, he at first felt anxious.

The new owners might want to fire him, to bring in someone younger. Instead, he found he was to keep his old position, apparently indefinitely.

Naturally, he would have to adapt to the new style of running the firm. Several of the executives were invited to France, to meet the new owners.

They would visit their headquarters, and see how they liked to run things.

The trip was organised on an almost luxurious scale.

They flew from London to Paris, and limousines met them at the airport. Plamer had sat next to Hugh Durnford, who had once been to France on holiday.

He was a great bore, and spent the whole trip talking about golf.

The hotel they were to stay at was obviously an excellent one.

A representative of the French company welcomed them there.

Clearly, they new owners wanted to make a good impression.

Palmer was certainly impressed when he saw their business methods.

Despite his long service, he always responded to new management techniques.

This may have been one reason for keeping him on after the take-over.

From what he saw, it was clear the future held some exciting changes.

- 1) Who spent the whole trip talking about golf?
- 2) Where was the French company based?
- 3) Why was Palmer anxious after the take-over?

style

(14) Mr. Turner had opened his barber shop on the corner thirty years ago.

In those days there was little competition, and he did good trade.

He never tried to learn any fancy skills, but gave good plain haircuts.

He was proud that his only tools were a comb and a pair of scissors.

Recently, disillusion had set in, as he saw that this was no longer enough.

Most of his customers these days were young schoolboys or retired old men.

The young men and teenagers preferred to go to new places in the town centre.

Trade was no longer so good, and Mr. Turner often had money worries.

He had passed retiring age himself, but felt that he had to keep working.

Besides, his shop had become almost a club, where the old fellows met.

Every Saturday morning the regular crowd turned up at the usual time.

Though supposedly there for a trim, they really came to talk and be warm.

Jack and Cyril would play a game of dominoes on a small table.

The others sat by the walls, reading old magazines as they

waited their turn.

Mr. Turner had kept his prices low, in order not to drive them away?

His kind, old-fashioned ways were the main cause of his lack of success.

- 1) How long had Mr. Turner had the barber shop?
- 2) Why had he kept his prices low?
- 3) Where did the young men now go to get their hair cut?

trim

C. DISTRACTOR PASSAGES

- (1) There were surprisingly few people on the golf course that Sunday.

When Barry drove into the car park, for once he found a place easily.

The slight rain that was falling had probably put people off.

Trevor, Doug, and Roger were waiting for him, and they soon began to play.

They played a foursome together regularly, every Sunday morning.

The drizzle was only a minor nuisance, and soon Barry and Doug were leading.

Over the last nine holes the others fought back, and eventually won.

Sitting in the clubhouse afterwards, the losers had to buy the drinks.

This was only a token gamble, because they mainly played for fun.

None of them were especially good players, but they still enjoyed it.

Normally, they stayed in the bar for an hour or so after the round.

This Sunday was different, because Roger had to leave early.

The others decided that they might as well go too.

He was planning to do some decorating that afternoon.

- 1) Was it usually hard to find a place in the car park?
- 2) What was the token gamble staked on the game?
- 3) Why did Roger have to leave early?

People had probably been put off by the slight rain that was falling.

wasp

- (2) On Wednesday morning Bob had to get up unusually early.

That was the day he had to sign for his unemployment benefit.

Every fortnight, people out of work sign a claim form to get their money.

When he got to the office, there were already many people waiting.

He had been unemployed for six months, and the queue kept getting longer.

Because everyone had a regular time to sign, he got to know the faces. Several of his friends were also out of work, and signed on the same day.

Bpb stood in the queue, talking to Mick and Ronnie, in front of him.

Bruce, further up the queue, had brought along a small white terrier.

Some people might not approve of bringing dogs to a place like that.

In a confined indoor space, animals could make a nuisance of themselves.

He was, however, considerate enough to keep it on a lead.

The queue was moving very slowly, and people were getting annoyed.

Normally two clerks were behind the counter, but today there was only one.

A flu virus was going the rounds, and there were lots of absentees.

When at last Bob reached the counter, he signed the form quickly.

Even with nothing to do all day, he resented having to queue for so long.

- 1) What day of the week was it?
- 2) Why was there only one clerk behind the counter?
- 3) How long had Bob been without a job?

When he got to the office, there were already many people waiting.

counter

- (3) At eight o'clock, right on time, the coach pulled into the bus station.

The waiting people picked up their baggage, and started to get on.

The coach was for a special outing, taking them to a day at the races. Roy and Joe sat near the middle of the coach, with Gary in front of them.

He had to lean back over the seats to join in the chat.

They spent the journey going over the runners in that morning's paper. When the coach got to the racecourse, they had to pay extra to get in. The coach ticket only covered the journey, not the course entrance money.

They had arrived in good time, and the first race would not be for a while.

They found a good spot to watch from, in the stands by the finishing post.

Gary had brought a pair of binoculars, and let the others borrow them. Before each race, they went down to place their bets.

Listening to the odds and watching the tic-tac men was part of the fun.

There were some good races that day, with exciting finishes.

Mid-way through the afternoon they bought some pork pies to eat.

Occasionally they went down to the bar to get some beer as well. Trudging back to the coach, they felt they had had a good day. Though they had lost money betting, that did not really matter.

- 1) At what time did the coach arrive at the bus station?
- 2) What did the three friends have to eat?
- 3) Did they win any money by betting?

The coach ticket covered the course entrance money, as well as the journey.

baggage

- (4) On Bank Holiday Monday, the Carters went for a day trip to the seaside.

They set off early in the morning, with the sun only just rising. It was a three-hour drive, and they hoped to miss the traffic jams. The car was full of picnic items, folding chairs, and things for the beach.

At first the driving was quite pleasant, but near the coast it was busier.

They stopped once for a few minutes, to stretch their legs.

When they reached the sea, they all went onto the beach.

The two children ran straight down to the water, to go paddling.

Mrs. Carter went with them, while her husband set up the chairs.

For lunch they had sandwiches and a flask of coffee, with fruit to follow.

The children made some sandcastles, building as high as they could.

Later, they all took a stroll along the promenade.

They paid a visit to the funfair, and went on some of the rides there.

Everyone had a marvellous time, especially with the sun shining.

On the way home, the children fell asleep on the back seat.

- 1) Why did the Carters set out so early?
- 2) What did they have to drink at lunchtime?
- 3) What was the first thing the children did at the beach?

apple

- (5) On Friday night, as usual, Len and Terry went out for a drink.

They went first to the Red Lion, near to where they lived.

There was a cold nip in the air, and they were glad to get inside.

It was still early, so at first they sat down to drink, talking.

Later, as crowds built up, they moved and stood with some friends.

Bernard was there, and he suggested going to another pub.

They walked down into the town centre, and had one at the George.

The jukebox was playing too noisily, so they moved next to the New Inn.

This was really their favourite pub, and they staued for a while.

Bernard began to get quite merry, and took to drinking vodka.

The others stuck to their usual pints, but laughed at his antics.

As he got noisier, it became a little embarrassing.

He did have a habit of becoming troublesome, if he drank too much.

The landlord kept looking over to where they were sitting.

Luckily, the pub would soon be closing, and they could go home.

- 1) Why did they leave the George?
- 2) What did Bernard take to drinking?

3) Was it a warm night?

brandy

(6) Philip had applied for the job more or less as an act of desperation.

His paper qualifications were good, but he was unsure of his suitability.

He had applied because he had been unemployed since graduating.

Now, though, he had received a letter asking him to come for an interview.

This was encouraging, and he did his best to make a good impression.

He put on his suit, and wore a tie, for the first time in months.

His shoes were polished until he could see his face in them.

The interviews were held at the company headquarters, a tall office block.

Half a dozen other candidates were waiting when he arrived.

The room was hot and stuffy, with the sun shining full on the windows.

One by one, a secretary called each applicant and led them out.

When Philip's turn came, he followed her nervously, tugging at his jacket.

There were three people sitting behind a table, with a chair in front of it.

Philip sat down, and they at once began asking him questions.

It was a tough half-hour, and he found himself stuttering a little.

At the end, he had a chance to ask questions of his own, about the firm.

He did so half-heartedly, as he felt sure he would not get the job.

1) How many people were there on the interview panel?

2) What did Philip wear to make a good impression?

3) Where were the interviews held?

The room was hot and stuffy, with the sun shining full on the windows.

stuffy

(7) On Saturday. Tim went for a stroll through the local street market.

He liked the bustle and sense of life that filled the place.

There was always something to see and hear, if you had the time.

He felt sorry for harried shoppers, rushing to get their weekend supplies.

They were unable to enjoy the market's essential charm and atmosphere.

Tim stopped for a moment to listen to a salesman's stream of patter.

The man claimed to be offering tea sets at a bargain price.

He was very persuasive, trying to get people to buy the goods.

Moving on a little way, Tim came to an antiques stall.

This was one of his favourite places, and he stopped to look at the items.

He loved old things, liking to admire them even when unable to buy.

Of course, most of the stuff was junk, but there were a few nice pieces.

He asked the stallholder the price of a few things, just to be polite.

Further along the row of stalls, Tim found someone selling second-hand books.

There were a lot of ragged paperbacks, but also some very fine hardbacks.

It was a volume of Dickens, beautifully bound, and only a little dog-eared.

He always bought at least one thing when he visited the market.

1) What were harried shoppers unable to enjoy in the market?

2) Did Tim buy anything at the antiques stall?

3) What was being offered at a bargain price?

Of course, some of the stuff was junk, but there were a lot of nice pieces.

stallholder

(8) Suddenly there was a great roar of triumph from the North Road stands.

In the closing minutes of the match, the home team had at least scored a goal.

When the whistle blew at the end of the game, they were still in the lead.

The cheering and shouting from the local supporters was deafening.

As the football ground emptied, the police kept the rival fans apart.

Leo and Brian made straight for the car park, behind the terraces.

Trevor had offered to drive them home, and he was waiting by his car.

He drove quickly out onto the road, anxious to avoid the rush of traffic.

The three of them talked about the match, and about the club's prospects.

They lived in a small town a few miles away, and were loyal fans.

Because they were feeling hungry, they pulled over at a fish and chip shop.

They did not like the pies and hot dogs that were sold at the ground.

Now, with the new alcohol rules, it was even impossible to get a drink there.

1) Was it the home team or the visiting team that won the match?

2) Where was the car park?

3) What food was sold at the football ground?

strike

(9) At weekends, Kevin often went out for a meal in a restaurant.

He did not enjoy cooking, and liked to have at least one decent meal a week.

On Sunday, as the weather was fine, he set off to walk into town.

He stopped to buy a newspaper on the way, to read while he was eating.

At the restaurant, he luckily found his regular table was free.

Sunday lunchtime was often busy, and he did not like sharing a table.

The waiter knew him, and soon brought him a menu.

After giving his order, Kevin sat back and flicked through the newspaper.

Occasionally he looked out of the window, at the people passing by.

Service in that restaurant was always very good, and his meal soon arrived.

Afterwards, he had a cup of coffee, and finished reading the paper.

Then he paid his bill and left, leaving a small tip for the waiter.

The sky had clouded over, so he walked home quickly to avoid any rain.

- 1) What did Kevin buy on his way to the restaurant?
- 2) Why did he walk home quickly?
- 3) Did he have to share a table that day?

Service in that restaurant was always very good, and his meal arrived soon.

waiter

(10) For weeks, there had been growing fears of a possible transport strike. The men had asked for a large pay rise, which the employers had refused.

Mark did not have a car, and relied on the bus service to get to work. He had taken his driving test three times, and failed on each occasion. On the day the strike was finally called, he faced a four mile walk. Luckily it was summer, and so he decided he might as well try it. He set off early, to be sure of arriving on time, carrying his briefcase. Most of the way he walked on the main road, which was very busy indeed. People who normally took the bus had today been forced to drive. Unusually, many of the cars had three or even four passengers. There were a lot of people on bicycles, and just as many walking. Mark had been walking for about half an hour when a car pulled up beside him.

It was a colleague from work, who offered him a lift.

They arranged that he would pick Mark up each morning from then on.

- 1) How many times had Mark failed his driving test?
- 2) How long had Mark been walking, before getting a lift?
- 3) What was he carrying?

conductor

APPENDIX 4

READING SPEED PASSAGES - EXPERIMENTS 2-6

Each passage is followed by the single recognition test word
presented in Experiment 4

(1) St. Joseph's Primary School was situated on the outskirts of Renthorpe.

As it was a Roman Catholic school, it drew its pupils from all areas of the town. Some of the children even had to make two bus journeys to get there. In the summer, this was never a problem, but in winter there could be difficulties. If the weather was bad, the long trip could be unpleasant, and sometimes parents would refuse to let their children set out. So far this winter, the weather had been relatively mild, but today the headmaster, Peter Simmonds, was getting worried. The forecast the previous night had been bad, and in the morning it was very cold. He stood at his office window, looking up into the solid grey ceiling of clouds. There was definitely snow in them. Towards the end of the mid-morning break, the snow began to fall. It fell gently at first, with huge white flakes blowing in the slight wind. The children loved it, of course, and were disappointed when the teachers called them in. Even as they were trooping back into the school, the snow began falling thicker and faster. In a few minutes, it had turned into a blizzard. The wind picked up, and as the snow began to settle deep drifts soon formed against walls and hedges. After an hour, there was no sign of it slackening off, and Simmonds began telephoning the bus depot. He soon learned that if the snow continued, the buses would not be able to get through. This had happened once before, and at tea-time he had been left with a school full of children unable to get home. After some more frantic telephoning, he made the decision to close the school early, to be sure that no-one was left stranded. As the school was so far outside the town, he really had no other choice.

1) When did it begin snowing?

2) What religion was the school?

3) Who telephoned the bus depot?

forecast

(2) John's car pulled up outside Peter's house at six o'clock in the morning. Dave got out of the back and ran up to the front door, but before he could ring the bell Peter came out. He had been watching for them through the window. Like the others, he was wearing large walking boots, which had seen a lot of use, and had his trousers tucked into his socks. He was carrying a small rucksack, which he stowed in the boot of the car. Then they set off. They were driving out into the country, for a day's walking over the moors. The three of them often did this at weekends, and enjoyed getting out into the fresh air. They drove north, and soon there was no sign of the bleak, dirty, industrial city they had left behind. The summer sun was shining, and the fields and woodlands looked green and inviting. After about an hour, they reached a small village, and pulled into a lay-by. This was to be their starting point. Peter collected his rucksack and they walked out of the village up a gentle slope. In a few hundred yards they came to a stile, where they left the road and headed across country. They walked steadily, taking their time to look around them and enjoy the day. They were making for Bramstone Rocks, which was some distance away and very high up. When they got there, they decided the climb had been worthwhile. The view was tremendous. The whole county seemed to lie stretched out before them. John claimed that he could make out the thin line of the sea in the distance. They ate their sandwiches sitting on the edge of the cliff. They were reluctant to begin the walk back, but in the end Dave stood up and made a start. The

others scrambled after him. At least going down hill would be easier.

- 1) At what time in the morning did the three set off?
- 2) Which direction did they drive?
- 3) Why would the walk back be easier?

cliff

APPENDIX 5

READING SPAN TEST SENTENCES - EXPERIMENT 2

A. PRACTICE SETS

One could already distinguish the rumble of the wheels from the sound of the engine.

She had finally made up her mind to go home and face the music.

He would come into the bar and wait until he was included in a round.

His father, a wrinkled old man, lay dying in the impersonal hospital ward.

Lying near the door of the hut was a torn and faded carpet.

He sang his piece with great feeling and volume and was warmly welcomed by the crowd.

It is very difficult for someone to believe in both science and religion.

Some willow bushes were growing in a thick clump on the far side of the stream.

He thought she might reveal unknown depths of strangeness, though not, perhaps, of meaning.

A sentry was marching slowly up and down in front of the gate.

B. EXPERIMENTAL SETS

To my surprise, I had happened to meet my former wife, Lillian, in the spring.

Far down in the avenue she could hear the sound of a street organ.

He sat beside her, put the glasses on the floor, and started to open the bottle.

He rushed up the stairs, hurried across the footbridge, and just managed to catch the train.

You could tell that he had got on by his travelled air and his expensive suit.

He had allowed himself to drift into a frame of mind that was conventional and stupid.

He wore a filthy singlet and ragged trousers, with a battered hat on his head.

The mechanic examined him closely before resuming his conversation in a subdued voice.

Teddy was a stout soft man in loose grey clothes and brown canvas slippers.

The pine trees rose up steeply on all sides, standing out against the blue sky.

He was horribly shocked by the events, and yet he was not altogether surprised.

She was jealous of the women I knew, but also of my friends and my books.

Every open space in the city was crowded with young people engaged in violent activity.

The eyes were small, canny, and curious, set in a brown and mottled face. The next morning I went down to look at the little house in the nearby street.

A little winding path led them down the hill to the river bank.

The lorry driver sat in his shirt sleeves on a cool June day.

She wore some sort of uniform tunic, though with no badge or stripe of rank.

The feeble old horse dragged the overloaded cart along at a snail's pace.

It was a light sunny evening when he emerged into the quiet and empty road.

His loneliness was the more intense now because he was surrounded by an alien culture.

He was a poor butcher. and by buying bad meat he ruined his business.

The clouds now rose like mountains and the coast was only a long green line.

Each blue car received a round of welcome as it topped the crest of the hill.

She dipped the towel in the basin and refreshed her eyes with the cool water.

There were so many different moods and impressions that he wished to express in verse.

She brought out a decanter of sherry and invited us to take a little wine.

I sat down in the cool arcade that surrounded the square and ordered a beer.

They got into a rowboat at the slip and made out for the American's yacht.

People have been facing death since the beginning of history, including millions of young men.

Nearly all the stalls were closed and the greater part of the hall was in darkness.

At one end of the table, ready for carving, lay a joint of roast ham.

The curtains had been pulled together to protect the carpet from the August sun.

A man of that age, bald and fat, with spectacles, is disgusting when drunk.

I was the youngest boy in the form, but was generally top of the class.

The door gave easily enough and he walked up a dark flight of narrow wooden stairs.

The stars were shining in the indifferent sky as I walked back to my hotel.

It was upon making this discovery that she paid her first visit to the lawyer.

The train for Rome was due to leave at a very early hour in the morning.

The teacher waited under the oak tree until midnight, and then went home.

Like most people in the Far East he began his lunch with a glass of gin.

In the past, he had been on friendly terms with a great many eminent people.

The sun went in behind some clouds and it began to grow cold.

From the church steeple a long thin shadow lay across the empty town square.

When she opened the door her expression was dark and set but she looked composed.

After a light breakfast he would dress warmly and go for a walk around the park.

They looked vacantly after some figures in the crowd and sometimes made a critical remark.

He let his tongue run on, to stifle the feeling of heaviness in his heart.

The studio was large and bare, with several chairs and a grand piano. Night after night I passed the house and studied the lighted square of window.

He had lived forty years with her, but they had passed as though in a fog.

There were several large turtles in the lagoon, and many different types of fish.

Many of the fishermen made fun of the old man and he grew angry.

The next morning, at about half past nine, John rang for a taxi.

In the living room, her parent's bookcases were filled with classic novels and poetry.

In the mornings I was driven around on guided tours of the hilly city.

The court was a big square room with wooden benches on one side.

When the restraining influence of school was at a distance I began to long for escape.

It was an Elizabethan mansion standing in a spacious park, in

which there were fallow deer.

Behind the drawn curtains a bee could be heard, knocking against the window pane.

APPENDIX 6

EXPERIMENTAL PASSAGES - EXPERIMENTS 3-5

Where appropriate, each passage is followed by the New recognition test word presented in Experiment 5.

(1) Jack and Ruth had three daughters and two sons, all married with children.

At Christmas, the whole clan came to see them, and Ruth cooked a special meal.

She was an excellent cook, and her Christmas lunches were always marvellous.

By tradition, they all sat around the big table in order of their ages.

At the foot, and at a small side table, were the young grandchildren.

Graham sat opposite his wife Louise, about half-way along the table.

Then came Anna and her husband Clive, and Jane with Stuart.

Ellen and James came next, while Alan, the eldest son, was near the head.

Carol, his wife, sat opposite him, and would help by serving out the soup.

FOREGROUND SECTION

A huge covered tureen was placed in the middle of the table.

On one side lay a deep, long-handled ladle, and on the other a stack of bowls.

As each bowl was filled, it was passed along from hand to hand.

The soup steamed slightly, it was so piping hot.

There was a delicious smell as the ladle dipped into the creamy liquid.

It was important to serve quickly, so the soup would not get cold.

The bowls clattered against the ladle as they were filled.

There was just enough soup left to fill the last one.

BACKGROUND SECTION

Above the table, brightly coloured decorations dangled from the ceiling.

There were sprigs of holly in all the corners, and some above the window.

All the walls were covered in streamers and Christmas cards.

There were more cards standing on the sideboard and mantelpiece.

A large tree stood in one corner, decked in tinsel and lights.
 Underneath, there were many mysterious looking boxes and parcels.
 These were the presents, and would be opened later in the day.
 The children could hardly contain their excitement.

She was careful not to get any drips on the tablecloth.

- 1) Who was careful not to get any drips on the tablecloth?
- 2) How many children did Jack and Ruth have?
- 3) Where did the young grandchildren sit?

gift

- (2) It was Whit weekend, and the funfair had come to town again.

The southern half of the central park was covered in tents and stalls.
 On Friday night, some friends met at a cafe, and then went down to the park.
 When they got to the funfair, at first they all went around together.
 After a while, though, they decided to split up.
 Laura wanted to go on the dodgems, and so did Maureen and Jack.
 Susan went with Pete to the rifle range, while Clare went on the
 big dipper.

FOREGROUND SECTION

This was said to be the tallest roller coaster in the country.
 The posters all promised the most exciting ride of a lifetime.
 As the train slowly crawled up the first slope, the passengers held on
 tight.
 A sturdy iron bar across the lap made sure people were safe.

Although the ride was meant to thrill and to frighten, there was no danger.

The train paused at the top of the climb, before hurtling down the first dip.

Speed increased rapidly, and it seemed certain to go off the rails. Then with a jerk the track pulled the train sideways, up the next hill.

BACKGROUND SECTION

It was a warm, dry evening, and the sky was still quite bright. There were already lots of people milling around the stalls and sideshows.

The fair was always popular, and more crowds would come as time passed. It was a welcome attraction over the Bank Holiday weekend.

At this early hour, there were many young children, with their parents. Later in the evening, it would be mainly teenagers and older people. Though the crowds were usually well-behaved, there were sometimes fights.

This generally happened a little while after the pubs had emptied.

She shrieked in terror at every twist and sudden drop.

- 1) Who shrieked in terror?
- 2) Where did the friends meet that night?
- 3) What time of the year was it?

opera

(3) They spent most of that Saturday afternoon working in their large garden.

The house was new, and there was a lot for the young couple to do.

After a few hours, they came inside and Pat started to prepare a meal.

She was peeling some potatoes when the knife slipped unexpectedly.

There was little pain, but it still gashed her wrist quite badly.

Her husband came and helped her to wash the cut under the tap.

He fetched the first-aid kit, but the bleeding refused to stop.

In the end they became worried enough to telephone for an ambulance.

At the hospital, the ambulanceman helped her down out of the vehicle.

A porter held open the door, and inside a nurse asked what was wrong.

The doctor soon came and examined the injury.

FOREGROUND SECTION

The curtains of the small cubicle were pulled over, for privacy.

Then carefully the blood-soaked bandages were peeled aside.

This was done very gently, as they had stuck to the wound in places.

A request for a bowl of water was quickly obeyed.

With a piece of cotton wool, the dried blood was cleaned away.

This left the injury easier to inspect and to assess.

Weighing up the damage turned out to be a simple task.

Once the wound had been cleaned, the examination was completed in seconds.

BACKGROUND SECTION

The casualty ward was relatively quiet, with only a few people waiting.

Despite this, the staff still seemed to be fully occupied.

There was a football match that afternoon, and that could mean trouble.

If there was fighting, both policemen and fans would come for treatment.

Things had to be got ready, in case they were needed.

It was a pity that people seemed unwilling to just watch the game in peace.

Things were often just as bad after a demonstration or a march.

Any kind of public disorder always meant more work for the hospitals.

He could see that several stitches were going to be needed.

- 1) Who could see that several stitches would be needed.
- 2) Where were the couple working that afternoon?
- 3) What was Pat doing when she cut herself?

yellow

- (4) The furniture store on Bridge Street caught fire early on Tuesday morning. The blaze was discovered by the night watchman, who called the fire brigade. In minutes, the street was full of fire engines, police cars, and ambulances.
- The fire was very dangerous, because of fumes from the artificial fibres. An ambulanceman took care of the watchman, who had been affected by smoke. A bystander offered to help, but a policeman told him to keep clear. The manager of the store, who had been notified, arrived in a little while. A fireman began hacking at the side door with an axe, but made little headway.

FOREGROUND SECTION

The door was solid, thick, and made a formidable barrier.

The rain of heavy axe blows barely dented the hard wood.

Usually, it was easiest to chop away either the lock or the door hinges. This time, neither part of the door seemed likely to give way. Despite the heat and showers of sparks, the effort had to be kept up. The door had to be broken down in order to get into the building. As the axe swung against the door, it made a great thudding noise. There was now a small splintered hole near the door handle.

BACKGROUND SECTION

At the front of the building, flames were licking through the windows. Houses on the other side of the street were having to be evacuated. Small children, still asleep, were wrapped in blankets and carried away. No-one had time to save any of their valuables or other possessions. At each end of the street, barriers had been erected to keep back the crowds.

For some reason, a big fire will always attract lots of onlookers. The flashing blue lights of police cars played across their faces. Everyone was warmly dressed, because of the cold night air.

He stood back at last and shouted for a crowbar.

- 1) Who shouted for a crowbar?
- 2) Where was the store?
- 3) Why was the fire especially dangerous?

hose

(5) The cruise had been advertised as the dream voyage of a lifetime.

It would be a leisurely sail around the most beautiful Pacific islands.

At first, the month-long trip went well, living up to all expectations.

The ship was luxuriously fitted out, and the entertainments were excellent.

There was dancing every night, a casino, and regular cabaret shows.

They were sailing over a stretch of little-known ocean when disaster struck.

The ship was suddenly shaken by a collision with a submerged reef.

To keep people calm, and avoid panic, the crew told them that all was well.

In the lounge, the purser tried to keep the entertainments running.

The barman was told to give free drinks to everyone.

On deck, the mate went around telling people not to worry.

Despite this, the captain was soon forced to order abandon ship,

The chief engineer was working in the hold, trying to repair a massive leak.

FOREGROUND SECTION

Water was flooding in through the damaged side of the ship.

It was already three feet deep in the hold, making work very difficult.

The only hope was to jury-rig some kind of temporary repair.

Although the vessel could not be saved, the sinking could still be delayed.

The backbreaking effort was worthwhile, to let the passengers get away.

There could be no consideration of personal danger at a time like this.

Staying down there, despite the urge to run, was an act of real heroism.

BACKGROUND SECTION

When the alarm bells sounded, passengers ran to their lifeboat stations.

Parents searched frantically for their missing children.

There was shouting and screaming, and people were knocked over in the rush. A lot of people did not have lifejackets, and were very frightened. Luckily, the weather was good and the sea was calm. This meant that launching the boats would be relatively easy. Given reasonable luck, another ship would be passing nearby soon. There was always a quick response to a distress call from a passenger ship.

He was trying to secure a patch over a jagged hole.

- 1) Who was trying to patch a jagged hole?
- 2) How long was the cruise intended to last?
- 3) What did the ship collide with?

(6) Pelham Engineering Company had always had a good profit record.

Shares in the company were a sound investment, and the future looked bright. Then suddenly one year the firm recorded a disastrous loss. Share prices fell, and an emergency shareholders' meeting was called. They wanted an explanation for the unexpected change in fortunes. As they gathered in the meeting room, their mood was angry and confused. All the directors and senior executives sat behind a table on a raised stage. In the centre, the chairman was trying hard not to look worried. The company secretary talked nervously to the senior sales director. The firm's chief accountant was there, ready with the necessary figures. At last, the managing director stood up to make the opening speech.

FOREGROUND SECTION

This was a big challenge, and much depended on giving a relaxed performance. In an emergency like this, the first things said would be crucial.

The shareholders had to be calmed down, but a confident style of presentation. A self-assured air of command was needed to control the meeting. The situation called for a firm voice, with no trace of hesitancy or doubt. In the first minutes the manner of delivery mattered more than what was said. There must be no sign of anxiety, or of attempts at concealment. The address needed to be frank, open, and free of deceit.

BACKGROUND SECTION

Latecomers were still arriving, looking for empty seats. There had never been such a crowded meeting before, in all the firm's past. Usually, only the major shareholders would bother to turn up. It took a serious crisis to bring out all the others. The fear of losing money was always a perfect cure for apathy. Many of the newcomers did not quite know how to behave. They looked around them, trying to copy what the others were doing. Those with only small numbers of shares felt lost and out of place. He took hold of the microphone and began to speak.

- 1) Who took hold of the microphone?
- 2) Where did the directors sit?
- 3) What type of industry was the company involved in?

book

(7) Stein's Menagerie toured Central Europe at about the turn of the century.

It had shows and stalls, but the main attraction was a zoo of predators.

Stein, the owner, knew his business, and kept only the fiercest carnivores.

Old Hans went ahead of the show, advertising the wild beasts to be seen there.

People in the small villages listened wide-eyed to the fearsome tales he told.

They would eagerly pay to see for themselves, and were never disappointed.

With its huge mane, the lion seemed strong and proud, truly a king of beasts.

The fine coats of the leopard and cheetah seemed to conceal their ferocity.

The South American puma, all in black, had an air of mystery.

FOREGROUND SECTION

Anyone looking into the cage would feel their hair standing on end.

An almost physical wave of horror beat out from the dark, brooding lair.

There were no roars or growls, only slow, cold, overpowering hate.

The feeling of danger, of great peril, had a supernatural intensity.

Sometimes women were overcome by this great sense of dread, and fainted.

Even brave men would quail into silence as the fear crept up on them.

No-one dared to get too close to the cage, or to go there alone.

This awesome, forbidding menace was uncanny.

BACKGROUND SECTION

The tents and sideshows made a colourful, gaudy spectacle.

Cries and shouts from the showmen rose above the noises of the crowd.

There were all kinds of stalls for people to browse around.

They could play games of chance, watch jugglers, or visit the palmist.

This all cost money, naturally, but people liked to have a fling sometimes.

When the local money ran out, the show would move on somewhere else. The travelling life was a good one, despite the cold and discomfort. Setting up the show every time was hard work, but was still worthwhile.

The villagers felt sure that it was an evil animal.

- 1) Which animal did the villagers feel sure was evil?
- 2) What sort of stories did Old Hans tell?
- 3) Where was the menagerie touring?

circus

- (8) In Rayner Road, Sunday morning seemed to be the time for doing odd jobs. Frank, at number 13, was carefully filling some cracks in the front steps. Next door, Jeff was busily painting the ground-floor window frames. Luckily, it had turned out to be an ideal day, dry and with no wind. Further up the street, John was working in the garden.

FOREGROUND SECTION

It was a big patch of land, with over half an acre to be tended. The biggest piece of garden lay to the rear of the house. As spring advanced, the grass was getting longer, and had to be mown. The flower beds would have to be spaded over, ready for planting. A recent storm had damaged the greenhouse, and some panes needed replacing. Probably the toughest job would be to prepare the large vegetable patch. There were some small trees, too, which should probably be pruned. All these things would probably take several hours to complete.

BACKGROUND SECTION

Away in the distance, the sound of church bells could faintly be heard. There was hardly another sign of life on that otherwise quiet morning. Rayner Road was situated in a pleasant area, out on the edge of town. There were no main roads to bring traffic noise, and no unsightly factories. The houses were large detached buildings, each designed differently. People felt it was a mark of success if they could move to live there. House prices were high, but that simply added to the attraction. There is a little of the snob in all of us.

He went to the shed to fetch a digging fork.

- 1) Who went to fetch a digging fork?
- 2) What could be heard in the distance?
- 3) Why was it an ideal day?

Helen

- (9) The little village of Calford nestled in the heart of the countryside. The town dated back six centuries, and many of the old buildings still stood. They have been converted to modern use, but the original structures remained.
- The main buildings were all grouped around the town square. There was a fine old church, still used as a place of worship. On the other side was the manor house, which was now a popular hotel. The rectory was now a private home, and the granary was used for concerts.

FOREGROUND SECTION

The village held a music festival every year, in the summer.

Some of the best musicians and orchestras in the country would come to play.

The standard of the performances was always excellent.

In the evenings, the music would float out over the peaceful countryside.

Everyone who came to play seemed to enjoy the experience tremendously.

The audiences, too, always seemed well pleased.

They could sit back in comfort, and listen to the good music.

All the usual facilities were provided, and no faults could be found.

BACKGROUND SECTION

The village was listed in many tourist guides and handbooks.

Coach tours often came to visit, and admire the well-preserved past.

Tea shops had been opened, to cater for the visiting tourists.

These were well hidden, in order not to spoil the town's historic appearance.

The town was quite secluded, but there were main roads not far off.

These did not spoil the setting, but did allow relatively easy access.

In the winter, of course, the town was largely left to itself.

Only a very few keen people would ever come to visit then.

It had excellent acoustics, and had been carefully converted.

- 1) Which of the old buildings had excellent acoustics?
- 2) How long had the town been in existence?

3) What was the village called?

(10) The quietest section of the carpet mill was the finishing department.

Most of the work in that area was done by hand, mainly by women.

In the distance, the roar of the heavy looms could barely be heard.

The women sat around tables mending small flaws in rugs.

It was delicate, precise work, but after a while it became boring.

Sometimes the radio would be switched on to help pass the time.

The best distraction, of course, was being able to talk to the others.

Hilda, Christine, and Bessie always sat at the same bench.

Margaret, another of the group of friends, would usually be there too.

Jane was the youngest girl, and always had to make the tea.

FOREGROUND SECTION

There was a little kitchen over in one corner of the big workroom.

It had a small second-hand gas stove, with just two rings on top.

The milk was kept in a bucket of cold water, as there was no fridge.

There was no teapot, so the boiling water was poured straight into the mugs.

It was difficult to remember which mug to put sugar in, and which not.

Instead of a tray, there was the lid of a packing case to carry the tea on.

Afterwards, the washing up had to be done, which was more of a nuisance.

Still, it was better than having to get straight back to work.

BACKGROUND SECTION

The workroom was large and brightly lit, with windows on three sides. With such detailed, painstaking tasks to do, good light was essential. In winter, with the dark mornings, huge strip lights were turned on. It was a cosy, pleasant sort of place, especially with rain or snow outside. During the summer, the windows were flung wide, to let in the fresh air. From this high up in the mill, there was a beautiful view. The building stood by a fast stream, which had originally been used for power. There were open moors and fields rising up on each side of the valley.

She would put the kettle on five minutes before the break.

- 1) Who would put the kettle on?
- 2) Which department did the women work in?
- 3) What was the best distraction from the boring work?

park

- (11) Long ago, a great King fell ill, with a strange and unknown sickness. The chancellor sent a knight to fetch healing water from a magic well. This well was said to be hidden far away, beyond a huge forest. In the woods, an evil demon caused the knight to lose his way. Starvation seemed the likely outcome, and wild animals were all around. Then an old man came walking by, carrying a large sack. This was a wizard, who offered to help on the mission of mercy.

FOREGROUND SECTION

A simple incantation, a few magic words, should be enough to do the trick.
In the sack, there were many powerful charms and talismans.
Though heavy to carry, they were the source of all great magic.
Preparations for the act of sorcery were quickly made.
Carefully, a mystic design was scratched on the earth.
Two spoonfuls of an evil-smelling brown powder were sprinkled all around.
Next, a small gold image had to be placed in the centre of the pattern.
Finally, everything was ready, and the magic could begin.

BACKGROUND SECTION

There had been such a long delay, that the quest already seemed hopeless.
Even if the life-saving well was found, it might still be too late.
The sickness had been worsening, even before the search began.
There was no knowing how long it would take to cause death.
If that happened, the whole country would be thrown into grief.
The black flags would fly, and there would be no more merry-making.
The people would go into great mourning, for a full year.
It would be a disaster for noble and commoner alike.

He recited a spell, and there was a flash of light.

- 1) Who recited a spell?
- 2) Where did the knight lose his way?
- 3) Who sent the knight on his quest?

(12) Teller and Lane, Ltd., was a small subsidiary of a national insurance company.

They were usually left to themselves, and went their own way about things.

The staff took life very easy, as there was little supervision.

Then one day head office decided to send someone to inspect the place.

The day before, everyone did their best to prepare a good impression.

The cleaning ladies were told to make sure that everything was spotless.

Mrs. Goves, who handled vehicle insurance, started getting out the ledgers.

In the entrance hall, the receptionist even watered the potted plants.

The secretary was trying to tidy up the filing cabinets.

FOREGROUND SECTION

The drawers were in a hopelessly confused and jumbled mess.

Files of different colours had all been thrown in together.

There had been no attempt to keep the items in alphabetical order.

Documents had been taken out, and then put back wherever there was space.

There was supposed to be a system, but it was always ignored.

The whole tangled muddle was a tribute to people's carelessness.

It could all be avoided if folders were always put back where they were found.

No-one seemed to care that things often got lost in all the disorder.

BACKGROUND SECTION

A surprise inspection like this could mean trouble for the firm.

If head office was displeased, the place could be closed down.

The parent company had a ruthless business reputation.

Any decision made on the basis of this inspection would be irrevocable.

It would be relatively easy to transfer the work elsewhere.

It was unlikely that the operation would be only trimmed down a little. The only alternatives were continuing on the present scale, or total closure.

The place would not be viable at all at any less than its present size.

She began to sort out the papers, one by one.

- 1) Who began to sort out the papers?
- 2) Where were the potted plants?
- 3) Why had life been easy for the staff?

redundant

(13) The last big rally meeting of the year drew a large crowd.

With several top teams competing, it promised to be a good race.

The event was sponsored by a well known hi-fi manufacturer.

For the first few laps, the cars jostled for position.

The Fiat and the Volvo were racing almost bumper to bumper.

Slowly, the Peugeot crept up the field,

Soon, though, the Ford had managed to take an unbeatable lead.

FOREGROUND SECTION

Getting in front quickly was the key to victory, on that track.

With a reasonable lead, it was almost impossible to be overtaken.

There were very few places which offered even a chance of passing.

Staying in front would be no trouble, unless there was bad luck.

The biggest danger was of getting too confident, and being careless.

With no challengers in sight, there was a temptation to relax.
 This would increase the risk of coming off the track.
 So long as an accident was avoided, though, there was no problem.

BACKGROUND SECTION

The spectators lined the side of the track, packed closely together.
 Sometimes they came dangerously near to spilling onto the track itself.
 It was always this way when a race attracted large numbers.
 Despite the stewards' efforts, people could not be kept in place.
 One day, there was going to be a terrible accident.
 Down in the refreshments tent, the var was doing good trade.
 It was a cold day, and the hot pies were selling well too.
 Some people seemed to come just to spend their time eating and drinking.

It cruised past the finishing line over two minutes in front.

- 1) Which car passed the finishing line over two minutes in front?
- 2) Who sponsored the race meeting?
- 3) Would there be any more races that season?

flag

- (14) On the day of the local league snooker final, Frank Worth was taken ill.
 He had been rushed to hospital with serious stomach pains.
 Harry, Don, Ian, and the other members of his team were frantic.
 They discussed what to do as they waited for the taxi to the match.
 Sam Brayshaw, the league chairman, ruled they could pick a substitute.
 Only Jeff, a total novice, was available at such short notice.

FOREGROUND SECTION

This dealt a savage blow to the team's hopes of success.
With so much at stake, any weakness could be disastrous.
The opposing side, on hearing the news, became very self-confident.
Under the circumstances, the match would probably be a walkover.
Although some advice or hints could be given, that would not help much.
Snooker is not an easy game, and it takes time to acquire real skill.
The only thing to do was to make the best of a bad job, and play on.
The match could still be fun, even if the result now seemed predictable.

BACKGROUND SECTION

Quite a large crowd had gathered in the hall to watch the match.
Most of the spectators were men, as usual, but there were some women too.
In some of the clubs, women were at last being allowed to play.
They were not as good as the men yet, but they were improving fast.
All the seats had filled rapidly, and now people were standing in the aisles.
Plumes of smoke from cigarettes rose and collected near the high ceiling.
The atmosphere was stifling, and the room grew hotter and hotter.
People who had brought their coats in took them off because of the heat.

He surprised everyone by playing a lucky winning game.

- 1) Who played a lucky winning game?
- 2) What was wrong with Frank Worth?
- 3) How did the team get to the match?

(15) The dog show was gradually reaching its final climax.

All week, the judges had been selecting the best dogs in each breed. Now, the top prize was to be given to the best dog in the entire show. This was everyone's ambition, and the owners made careful preparations. In the kennels behind the show area, animals were given a last grooming. The little terrier was being carefully brushed for the judging. The poodle, the bulldog, and the Afghan were also being got ready. It was obvious that the red setter had all the winning qualities.

FOREGROUND SECTION

Anyone with experience could easily pick out the good signs. For competition at this level, all the features needed to be perfect. There was no doubt that the judges would look for all-round excellence. Good looks, pedigree, and obedience were just what was wanted. Overall healthiness would be another way to get good marks. Clean and even teeth, in perfect condition, would also be noticed. An especially important point was to have clear, well-shaped eyes. A good colour to the coat, with a little shine, would clinch the victory.

BACKGROUND SECTION

Out in the parade ring, everything was being arranged for the contest. Fresh sawdust was being scattered along the ground. A large table, for the judges, was placed to one side of the ring. A lane was marked out in front, for the competitors to be led along. Overhead, the bright lights from the ceiling cast many shadows. The spectators stood or sat, waiting expectantly for the start. Prize money this year was quite large, so there was much excitement. A loud hum of conversation throbbed around the hall.

It was the only dog which really stood a chance.

- 1) Which was the only dog that really stood a chance?
- 2) How long had the show been running?
- 3) Where were the animals given their final grooming?

cup

(16) Variety shows are one of the few thriving forms of theatre entertainment.

They have something for everyone, which ensures their popularity.

Seat prices are usually low, and the night out can be cheap.

Mary and Jim often liked to spend an evening out in this way.

The most recent show they had seen had really been very good.

The first performer was a comedian, who soon livened up the audience.

After that came a conjuror, and then there was a singer.

The ventriloquist was by far the most popular act of the evening.

FOREGROUND SECTION

Several of the dummies used were funny enough by themselves.

A good script backed up the successful visual impact.

The only other props were a chair and a large wooden trunk.

One by one, the dummies demanded to be let out of the chest.

Each one had a different personality, and a different voice to go with it.

The conversations were incredibly quick, and very convincing indeed.

All the voices came out loud and clear, and could easily be heard.

There were no slurred consonants, or awkward pauses and accents.

BACKGROUND SECTION

The small theatre was very full for this particular show.
 Advertisements in the local newspapers had drawn a good crowd.
 The show was solidly booked for the rest of the week, as well.
 Only by attracting the customers could such a place hope to survive.
 It was a nice, comfortable theatre, with soft seating at all prices.
 There was a bar upstairs, which served drinks during the interval.
 The same building housed an ice rink and the central library.
 It was the entertainments centre for a large area of the city.

He was given the longest applause of the evening.

- 1) Who was given the longest applause of the evening?
- 2) Was a night out at a variety show expensive?
- 3) Why were variety shows generally so popular?

balloon

- (17) A little row of shops stood to one side of Stanley Square, near the church.
 It was useful for the neighbourhood, as the town centre was two miles away.
 Instead of going so far, they could get things just around the corner.
 Then the council declared that the two of buildings had to be demolished.
 The shopkeepers called a meeting, to see what could be done.
 The grocer suggested that they should send in a petition.
 The newsagent and barber agreed, and so did the wine merchant.
 The butcher was more interested in getting compensation, and relocation help.

FOREGROUND SECTION

After all, there are some businesses which do better in the town centre.

The local custom was all right, but more meat could be sold in town.

This view was not greediness, but simple realism.

It was difficult to make a living selling only small pork pies and pasties.

Even at weekends, there were few people buying joints or large cuts.

The most expensive things anyone seemed to buy were two pounds of chops.

The cold store was never full, and nor for that matter was the till.

Most days, the delivery van brought no more than a few sausages.

BACKGROUND SECTION

If these buildings were knocked down, the area would lose a great deal.

Opposition to the council's plan was fairly strong in the locality.

It was not at all obvious just why the row was going to be demolished.

The buildings were structurally sound, and not hazardous in any way.

One theory was that some new development would be built there.

This would probably be a block of old people's flatlets.

The planners did not seem to care about damaging the existing community.

There was a lot of resentment in the neighbourhood.

He wanted to move to a place where trade was better.

- 1) Who wanted to move to a place where trade was better?
- 2) How far was the area from the town centre?
- 3) What building was near the shops?

licence

(18) The town hall housed not only the council offices, but also the county court. An usher directed a man to the courtroom, where a case was under way. The trial had paused, as the admissibility of some evidence was challenged. This could clear the accused, but the prosecution wanted it disallowed. The prosecutor stated his arguments precisely, in a superior tone of voice. The defending barrister was infuriated, and replied angrily and to the point. While this went on, the defendant sat quietly, apparently unconcerned. A reporter from the local newspaper took notes of what was being said. After a time, the judge asked the jury to leave the courtroom.

FOREGROUND SECTION

Such a request can be made for many reasons during a trial. In this instance, it was so the jury would not hear the item under debate. Before the jury heard the evidence, it had to be decided whether to admit it. The order to leave was obviously necessary, at that stage of the argument. The jury were told why they were being asked to leave. This ruling would allow freer discussion of the point at issue. When the debate over the point of law was completed, they could return. However, they had to be excluded while the discussions went on.

BACKGROUND SECTION

It was a hot, stifling day, and the court was stuffy and unpleasant. The case was a difficult one, and seemed likely to last a long time. In such weather, it was hard to concentrate or stay alert. There were few spectators sitting up in the public gallery. Usually, there were quite a few curious people up there. There would sometimes be school groups, there to see how courts worked.

On a day like this, though, most people had stayed away.

It was much nicer to be out in the fresh air and the sunshire.

His instructions were quickly obeyed by the jury members.

- 1) Whose instructions were quickly obeyed?
- 2) What infuriated the defending barrister?
- 3) Where was the county court located?

journalist

(19) It was the last day on location, and a crucial scene was to be filmed.

The film was a week behind schedule, so everyone got ready early.

It was six in the morning when work began.

The sound engineer carefully tested the microphones and recording set-up.

The cameraman and lighting man discussed a problem about some shadows.

The director gave the final instructions, using a large megaphone.

FOREGROUND SECTION

The loud noise of the amplifier made everyone stop and listen.

Clearly and concisely, the way the scene should be played was explained.

This vital scene had to be done perfectly, with everything just right.

It would be madness to trust people to know automatically what to do.

Even experts and professionals had to have all the details spelled out.

Every aspect of the filming needed close personal supervision.

All the requirements had to be made explicit before shooting began.

When people understood what was wanted, they could be left to get on with it.

BACKGROUND SECTION

When this shot was completed, the equipment would be dismantled. The entire crew would move to the studio, for the indoor scenes. There would be some regrets at leaving the outdoor location. One of the best features of the job was going to foreign, exotic places. It was usually possible to go sight-seeing, after working hours. Although living conditions could be primitive, it was still fun. Holidaymakers often paid a great deal to come to such places. Working in films often meant being paid to go there yourself.

He made sure the last orders were heard, then sat down.

- 1) Who made sure the last orders were heard?
- 2) How far behind schedule was the film?
- 3) What time did work begin?

stereo

(20) Near the end of Grove Street was a small block of private flats.

The red brick building stood discreetly to one side of the quiet road. Nearly all the people who lived in the flats were artists of some kind. On the ground floor lived a famous writer, next door to a painter. An actress and a soprano singer shared a flat on the floor above. A sculptress had a large flat at the very top of the building.

FOREGROUND SECTION

This flat was ideal, since it had a large studio attached to it.

It was light and spacious, and usually full of wood and stone.
A door led off from the living room into the studio itself.
The light came streaming through a wide bank of windows on one side.
Against one wall stood a bench, loaded with mallets and chisels.
In one corner there was even a small forge, for working in metal.
The room was well insulated, to deaden all the loud noises.
Sheets of paper, with preliminary sketches, lay all over the floor.

BACKGROUND SECTION

The building was owned by a company, which rented out the flats.
It was a very profitable enterprise, in that area.
Living accommodation was very scarce, and so rents could be high.
There was only a little maintenance to offset against the income.
The place was kept in good repair, and very little ever needed attention.
Perhaps the biggest item was cleaning, and also servicing the lifts.
There was an underground garage for residents to park their cars.
This could be locked up, protecting them from thieves and vandals.

She had had many pieces of work exhibited in local museums.

- 1) Who had had many pieces of work locally exhibited?
- 2) Where were the flats?
- 3) What was the building made of?

poet

APPENDIX 7

RECOGNITION TEST PRACTICE PASSAGES -

EXPERIMENT 5

- (1) On Saturday. Time went for a stroll through the local street market.
He liked the bustle and sense of life that filled the place.

market

- (2) Suddenly there was a great roar of triumph from the North Road stands.
In the closing minutes of the match, the jome team had at last scored a goal.
When the whistle bloew at the end of the game, they were still in the lead.

goal

- (3) Philip had applied for the job more or less as an act of desperation.
His paper qualifications were good, but he was unsure of his suitability.
He had applied becaise he had been unemployed since graduating.

post

- (4) Because of the bus strike, many peple had been forced to drive that day.
Unusually, many of the cars had three or even four passengers.
There were a lot of people on bicycles, and just as many walking.

bus

- (5) They were going down into town to see a new film at the cinema.

As they walked, they talked causally about work and people they knew there.

work

- (6) The place looked as though it has been hit by a bomb.

Plants had been torn up, greenhouses smashed, and sheds broken into.

Some tools had been stolen, but the general destruction was more serious.

vandal

- (7) Jerry went to the launderette very early that morning, as he always did.

He found doing his washing a nuisance, and liked to get it over with quickly.

Jerry

- (8) The Winston Players were a popular local amateur dramatic group.

There were always good crowds when they performed their shows.

crash

- (9) The library was a centre not just for book lending, but for many activities.

Children's story evenings, links with adult education, information services.

hotel

(10) Every Saturday morning the regular crowd turned up at the usual time.

Jack and Cyril would play a game of dominoes on a small table.

beer

(11) At weekends, Kevin often went out for a meal in a restaurant.

He did not enjoy cooking, and liked to have at least one decent meal a week.

eat

(12) Michael Palmer was a middle-rank executive in the firm of Garner and Sons.

He had worked up to that position over the years, and felt he did a good job.

When they were taken over by a larger French firm, he at first felt anxious.

executive

APPENDIX 8

EXPERIMENTAL PASSAGES - EXPERIMENTS 6-7

- (1) The hotel was of the highest quality, situated in the heart of London.
It was luxuriously fitted out, and was an expensive place to stay.
The excellent food was prepared by a team of French chefs.
As guests arrived, the commissionaire would hold open the doors.
At the reception desk, a clerk issued the room keys.
The porter had grey hair and wore a smart blue uniform.

FOREGROUND SECTION

There were brightly polished buttons down the front.
The jacket was trimmed with slightly darker material.
The hotel's name was stitched on the breast pocket.
The same symbol appeared on the front of the neat tie.

BACKGROUND SECTION

The lobby was large and spacious, with a high ceiling.
Armchairs were dotted around, screened by potted plants.
There were four lifts, including an express to the penthouse.
A small shop stood discreetly in one corner.

TARGETS

He looked much better than when in ordinary clothes.
He often thought of using a dye to look younger.

- 1) Where was the hotel?
- 2) What nationality were the chefs?

(2) Mrs. Greaves lived alone in a huge old house.

She had no family left, and so she took in lodgers.

The house was much too big for one person.

There were six rooms to let, and a wide variety of people stayed there.

Mark was a student, and so was Sheila on the ground floor,

Stephen was a music teacher in a local school.

On the top floor lived Graham, who worked in a bank.

FOREGROUND SECTION

From so high in the house, there was a marvellous view.

Birds nested in the eaves, and could be heard moving around.

Noises filtered up from the rooms below, muffled by the distance.

It was always cold in winter, living just under the roof?

BACKGROUND SECTION

The house stood in the middle of a large expanse of garden.

It was a good neighbourhood, out on the edge of town.

It was quiet and peaceful there, with little traffic.

The road had wide grass verges, and was lined with trees.

TARGETS

He disliked having so many stairs to climb.

He enjoyed helping people with their financial affairs.

1) What was the landlady's name?

2) How many rooms were there to let in the house?

(3) The regional orchestra was going on tour again.

Once a year, the company travelled all over the region.

They gave concerts in town halls and schools, as well as theatres.

The coach arrived at their headquarters at eight in the morning.

All the players were ready, with their instruments and suitcases.

The flautist and cellist helped to stow the baggage.

The trombonist had brought a pack of cards to use on the journey.

The violinist took the front seat, and read a newspaper.

FOREGROUND SECTION

This was really the best seat in the coach.

There was plenty of leg room, with no other seats in front.

It was nearest to the door, too, which was convenient.

Any trip is more pleasant in a comfortable seat.

BACKGROUND SECTION

These tours were a welcome change to the usual routine.

Throughout the region, audiences were enthusiastic and appreciative.

The travelling was an important part of the orchestra's role.

Many places would never see live music otherwise.

TARGETS

He liked being able to watch the road ahead.

He began with the sports pages.

1) How often did the orchestra go on tour?

2) What time did the coach arrive?

(4) It was just after eleven, and the coffee bar was crowded.

People always gathered there after having a class.

Because of the noise of the jukebox, it was almost impossible to talk.

Susan and Jill struggled through to the counter.

They got their coffee, and were lucky enough to find a free table.

They were joined by Mary, a mature student who never went to lectures.

FOREGROUND SECTION

Missing classes can be due to laziness, but it does have advantages.

Some people dislike being limited to the basic coursework.

To study only what is covered in lectures may seem boring.

It can be more productive to have a wilder spread of interests.

BACKGROUND SECTION

Over in the corner, the pinball machines were in constant use.

A darts match was also under way.

The dark, crowded room was stuffy and very hot.

Only when the door opened was there a breath of air.

TARGETS

She passed the exams only by solid private reading.

She had made few friends because of the age difference.

- 1) What time in the morning was it?
- 2) Why was talking almost impossible?

(5) The bus to work that morning was running half an hour late.

The weather was atrocious, with snow and freezing rain.

Traffic had crawled forward at a snail's pace.

Several girls from the same office sat together on the back seat.

They met on that trip every morning.

Jane had worked in the typing pool for six months.

Laura and Carol were punch operators in the computer room.

At the bus stop, Pat from the wages office slipped on some ice.

FOREGROUND SECTION

Working with the firm's cash was a very responsible position.

The money could be a great temptation for a greedy person.

Besides this, even an honest mistake could have serious consequences.

Every penny had to be fully accounted for.

BACKGROUND SECTION

The tall office block was brightly lit on that dark morning.

Overhead, the thick clouds hung low in the sky.

The weather forecast that morning had been for more snow.

It was the worst winter in living memory.

TARGETS

Her main job was to calculate the wage packets.

Her knee banged heavily against the kerb.

1) How late was the bus that morning?

2) Where did the girls sit?

(6) The football club had a glorious past, but now it was struggling. Winning the FA Cup three times in the twenties was no help at all. The stands were decaying for lack of maintenance. Usually, fewer than two thousand people came to the matches. The chairman had invested a lot of his own money in the club. After every defeat, the manager got the blame. The coach was on half-pay, and there was only one groundsman. The captain, a local player, was getting too old for the game.

FOREGROUND SECTION

As time passes, the joints stiffen and stamina ebbs away. Football is a tough sport, and it takes youthful fitness to survive. All the hard knocks had become more difficult to take. The training sessions alone were almost too much.

BACKGROUND SECTION

The ground was near to the centre of the town. For this reason, the property at least was quite valuable. There were some plans to develop a nightclub and restaurant. There was plenty of space to park cars.

TARGETS

He should have retired a long time ago.
He had been born in a village a few miles away.

- 1) How many times had the club won the FA Cup?
- 2) Why were the stands decaying?

- (7) It was almost lunchtime, and the restaurant staff were getting ready.
The dark, oak tables were being carefully laid.
Wednesday was always busy, and it was a market day.
Shoppers would come in loaded with bags to have a meal.
The manageress went round checking all the preparations.
In the kitchen, the cook was warming the ovens.
At the bar, the barmaid was polishing the glasses.
The waitress, who was new to the job, stood looking out of the window.

FOREGROUND SECTION

The glass was shaded by a white net curtain.
This made it possible to watch the street without being seen.
It was interesting to observe people who stopped to read the menu.
The cars and pedestrians were a constant distraction.

BACKGROUND SECTION

The doors would be opened at eleven thirty.
There were twenty tables in the downstairs dining room.
Upstairs was a private room used only for special functions.
It was especially popular for wedding receptions.

TARGETS

She liked to spend time just watching the world go by.
She had only worked there for just over a week.

- 1) Why was Wednesday always such a busy day?
- 2) What sort of wood were the tables made of?

(8) The unemployment benefit office was on a dingy side street.

A disused textile warehouse stood on the opposite corner.

Inside, the office was painted a drab, dirty grey.

It was a very depressing place .

On Monday, Anne went down to sign her claim form.

This had to be done once a fortnight.

Karen and Sue were already there, further up the queue.

The queue had reached all the way to the door.

Jenny arrived with her dog, and had to stand outside in the rain.

FOREGROUND SECTION

The big, friendly labrador refused to keep still.

It kept tugging on the lead and jumping into the air.

This lively nature was one of its nicest qualities.

A dog is usually an affectionate creature.

BACKGROUND SECTION

The prospects for any young unemployed person were very bleak.

All the traditional industries in the town were slowly decaying.

There were no new businesses being started.

It seemed as though the recession would go on for ever.

TARGETS

1) What colour was the benefit office painted?

2) What building stood on the opposite corner?

- (9) Early in the morning the town centre slowly began to stir.
- There were only a dozen shops in the small pedestrian arcade.
- By eight, all the shopkeepers were getting ready to open.
- As it was summer, the butcher let down the blind to shield the meat.
- The newsagent hung a rack of papers outside his door.
- In the sweet shop, the confectioner was polishing the counter.
- The grocer, who did excellent trade, came out to clean the windows.

FOREGROUND SECTION

The two large squares of plate glass were looking rather grimy.

The panes in the door, too, were in need of a wash.

The wet leather squeaked as it rubbed on the glass.

Little trickles of water ran slowly downwards.

BACKGROUND SECTION

The arcade was becoming busier, with people going to work.

There was a large office building at one end.

The town hall and county court were at the other side.

Not far off was the headquarters of the local building society.

TARGETS

He propped a ladder against the wall.

He was making a great deal of money.

- 1) How many shops were there in the arcade?
- 2) What season of the year was it?

(10) John and Cath took their children to the safari park.

It was the Bank Holiday, and so John was off work.

They drove through the tall gates of the park.

Luckily the sun was shining, making it a perfect day.

There were animals everywhere, all roaming freely.

A lion was sprawled fast asleep under a tree.

By a water tank, an elephant was taking a long drink.

There was a rhino, a giraffe, and a zebra.

A monkey, which was holding a banana, sat down in the road.

FOREGROUND SECTION

There was no room to drive around.

The cars were well and truly blocked from moving on.

The park rules forbade turning to go back.

All that could be done was to sit patiently and wait.

BACKGROUND SECTION

Safari parks are really an excellent innovation.

In some ways they are a lot better than zoos.

They have lots of space and open air.

In zoos, the cramped conditions can be very cruel.

TARGETS

It showed no inclination to get out of the way.

It took small bites every now and then.

- 1) Why was John off work?
- 2) What made it a perfect day?

(11) It was ten o'clock, and the board meeting was due to start.

The conference room at the top of the building was ready.

Blotters, pens, and notepads were neatly laid out.

The managing director was the first person to arrive.

Then came the treasurer and company secretary.

The sales director sent down for some coffee.

The chairman arrived late, carrying a heavy briefcase.

FOREGROUND SECTION

The traffic on the road to the office had been abominable.

A long tailback had built up, because of an accident.

The queue had moved forward painfully slowly.

No amount of horn-blowing had been of any help.

BACKGROUND SECTION

The view from the conference room windows was superb.

Like most office blocks, this one was extremely tall.

This great height allowed the entire city to be seen.

The streets below looked like a tiny model.

TARGETS

He had held up the meeting for half an hour.

He had brought several thick documents for discussion.

- 1) What time had the meeting been due to start?
- 2) What things were laid out ready?

(12) The morning routine in the hospital ward never seemed to vary.

At six, the sister in charge would start things off.

All the patients disliked being woken up so early.

An auxiliary would pull back the long green curtains.

By seven, a kitchen assistant had brought up the breakfast.

A nurse took the temperatures and tidied the beds.

FOREGROUND SECTION

The thermometers were kept in small tubes on the wall.

Each one had to be shaken down after use.

It was a relatively easy job, but demanded complete accuracy.

It was an important indicator of states of health.

BACKGROUND SECTION

This was always the noisiest time of day.

The morning chatter was loud and friendly.

Dishes clattered together and trays banged on tables.

The sound of a floor polisher dominated all the rest.

TARGETS

She wrote down the figures neatly.

She arranged all the pillows comfortably.

- 1) What time were the patients woken up?
- 2) What colour were the curtains?

(13) The main city theatre staged a variety show..that Christmas.

Pantomimes were old-fashioned, according to the new manager.

Luckily, press reports on the show turned out to be excellent.

After the last show, all the performers gathered in the pub.

The singer, who was the star, bought all the drinks.

The acrobats, a husband and wife team, had a new booking abroad.

The comedian and the juggler arrived later than the rest.

The magician drank tonic water and sat by the fire.

FOREGROUND SECTION

A huge hearth surrounded the open iron grate.

Bright yellow flames were licking up the chimney.

The coal was banked up very high.

The glow of the flames was reflected on nearby chairs.

BACKGROUND SECTION

The show had made a surprisingly big profit.

Takings were well up on the previous year.

After such a success, the experiment would no doubt be repeated.

The important thing in show business is to please the customers.

TARGETS

He could not stand the cold of winter.

He never touched any alcoholic liquor.

1) What did the new manager think of pantomimes?

2) Did the critics like the variety show?

(14) Due to a dock strike, the ship was stranded in a French port.

The cruise eventually had to be abandoned there.

Passengers were flown home at the company's expense.

As days dragged on, the crew found ways to pass the time.

The steward, purser, and engineer started a poker school.

The first mate took to fishing in the harbour.

The captain read a great deal and went sightseeing.

FOREGROUND SECTION

It was pleasant to sit back with a good book.

Reading is a relaxing way to spend free hours.

On top of that, there is a lot to be learned from books.

No-one has ever regretted time spent reading.

BACKGROUND SECTION

The strike, as usual, was over a pay increase.

The dockers wanted more than the authority could afford.

In France, such strikes are regular occurrences.

They seem to be even more frequent than in Britain.

TARGETS

He especially liked long historical novels.

He enjoyed looking around new places.

- 1) Where was the ship stranded?
- 2) How did the passengers get home?

(15) The annual golf tournament was a problem for the local police.

They had to control the thousands of incoming cars.

This year, a thick fog made matters even worse.

In the main car park, a Renault and a Ford collided.

Signs banning parking on the approach roads were often ignored.

A new Rover was found blocking a private access route.

The driver of a Fiat was prevented from parking nearby.

Later, a Mini was found on a sharp bend, with no lights despite the fog.

FOREGROUND SECTION

In the poor conditions, lights were really essential.

Even a little mist is troublesome, and heavy fog is worse.

The human eye cannot penetrate such a thick curtain of vapour.

Smaller vehicles seem to vanish altogether.

BACKGROUND SECTION

The tournament was due to start at ten o'clock.

All the players were gathering well in advance.

Last-minute preparations were complete, and everyone was ready.

There was a lot of prize money at stake that year.

TARGETS

It could hardly be seen, there by the roadside.

It was parked in a very dangerous place.

1) How often was the tournament held?

2) Where were the No Parking signs?

(16) The country court was housed in the local town hall.

A mural of industrial scenes decorated the outside wall.

Inside, the bustle of activity went on as usual.

A barrister walked by, talking to a solicitor.

An usher showed a journalist into the courtroom.

A witness waited impatiently to be called, sitting on a hard bench.

FOREGROUND SECTION

The case had gone on for much longer than expected.

A legal wrangle had held up the proceedings.

All the testimony should have been given earlier in the morning.

It was now afternoon, and the case was still dragging on.

BACKGROUND SECTION

The council offices were in the east wing of the building.

It was there that the business of local government was conducted.

Recent boundary changes had widened the authority's power.

It now covered a much bigger geographical area.

TARGETS

He was anxious about missing an appointment.

He felt very uncomfortable.

1) Where was the county court located?

2) How was the outside wall decorated?

(17) The small village of Barbridge lay deep in the Cotswolds.

Market fardening was the main source of income there.

Around the green, all the important buildings stood in a circle.

There was a shop, and a small one-pump garage.

The school church, and local hall stood nearby.

The pub, a very old building, was the real heart of the village.

FOREGROUND SECTION

A surprising number of valuable old properties still exist.

Built in an earlier age, they mean much more to us today.

He can learn a lot from seeing how our ancestors lived.

The past has a message for us all.

BACKGROUND SECTION

Horticulture was a profitable enterprise because of the climate.

Warm summers with reasonable rainfall are essential.

The soil quality also has to be very high.

With such conditions, plants will grow very easily.

TARGETS

It was listed as a building of historical interest.

It was the place where everyone gathered at night.

- 1) What was the village called?
- 2) What was the main source of income?

(18) The film was being shot in the far north of Scotland.

It was January, and a biting wind came down off the hills.

This was the last location work before moving into the studio.

The director was giving instructions to the cameraman.

Things had paused while the sound engineer checked the microphones.

A lighting technician adjusted some lamps.

An actor, who was wearing a fur coat, went for a cup of tea.

FOREGROUND SECTION

There was a refreshment wagon on the edge of the site.

It served all kinds of snacks and beverages.

The kettles were steaming on the electric stove.

A row of cups was lined up ready on the counter.

BACKGROUND SECTION

The film was a purely British production.

It had been financed mainly with money from grants.

Despite the low budget, it was an interesting project.

The cost of a film is no indication of its quality.

TARGETS

He was looking forward to having a warm drink.

He was snug and warm despite the weather,

1) Where was the film being shot?

2) Would there be any more location work?

(19) It was the last day of the race meeting.

The stables stood behind the saddling enclosure.

A stable lad led a horse around the yard.

As the owner looked on, a steward checked the list of runners.

The jockey was ready, but the animal seemed lame.

The trainer felt the foreleg, and muttered a few words.

FOREGROUND SECTION

The flesh was slightly swollen below the joint.

Compared to the other leg, it was quite noticeable.

The animal flinched a little as it was touched.

The injury was obviously painful.

BACKGROUND SECTION

The stands were crowded with eager spectators.

In the car park, coaches were jammed nose to tail.

The refreshment stands were doing good trade.

It was a hot day, and drinks were selling well.

TARGETS

He fingered the muscle as gently as possible.

He spoke too indistinctly to be heard.

1) Was it the first day of the race meeting?

2) Where were the stables?

(20) A group of friends sat around the living room, talking.

They had just got back from the local night classes.

Clare and Vera sat by the fire, as it was a cold December night.

Sally looked through the records on the stereo cabinet.

Pamela talked about a recent trip to France while pouring the coffee.

FOREGROUND SECTION

The steaming brown liquid gave off a delicious smell.

It was freshly made, and still piping hot.

A sugar bowl and jug of milk were placed on the table.

The cups rattled as they were filled.

BACKGROUND SECTION

Outside, the wind was starting to pick up.

It howled alarmingly around the house.

There was the sound of rain falling against the window.

It was certainly not a very pleasant night.

TARGETS

She was careful not to spill any drops on the carpet.

She had spent two whole weeks in Paris.

1) Where had the friends been that night?

2) What month was it?

(21) The express to King's Cross thundered along at a great rate.

After a delay at Peterborough, lost time had to be made up.

The driver kept the throttle wide open.

A buffet steward went around with tea and sandwiches.

The ticket collector made his way slowly down the train.

The guard was having a cigarette, sitting up in the baggage van.

FOREGROUND SECTION

This was a good place to sit in peace and quiet.

There were no chairs, of course, but that made no difference.

It was easy to improvise something from all the bags and suitcases.

A little imagination could soon work something out.

BACKGROUND SECTION

The train was passing over wide, level ground.

Further north, in the Pennines, the terrain was more hilly.

There were more tunnels and cuttings on that leg of the trip.

The southern and eastern counties were relatively flat.

TARGETS

He sat on a large trunk, feeling very comfortable.

He blew a series of neat smoke rings in the air.

- 1) Where was the train going?
- 2) Where had there been a delay?

(22) Colin Brown was a successful architect, and lived in a Georgian house.

He and his wife had seven children, so they needed a large home.

Of the five daughters, Alison was the quiet, bookish one.

Debbie and Louise, like the others, loved sports and outdoor activities.

Marion was an enthusiastic gymnast, winning many cups.

Helen was a red-haired girl, and enjoyed dinghy sailing.

FOREGROUND SECTION

The fiery mane had never been cut short.

It hung long past the shoulders, and bunched at the sides.

The blaze of colour instantly caught the attention.

It made for a very striking appearance.

BACKGROUND SECTION

Bringing up a large family can be very difficult.

In the early years, money had been very tight.

With so many mouths to feed, a guaranteed income is essential.

People who seem wealthy may not always have been that way.

TARGETS

She had a flaming temper to match the colouring.

She enjoyed the excitement of speeding across the water.

1) What was Colin Brown's profession?

2) How many children were there?

(23) The amateur operatic society was preparing another show.

The stage of the local civic hall was busy with activity.

Graham and Kathleen painted the final piece of scenery.

In the back rooms, Hilda and Hean made the refreshments.

Margaret always sewed the costumes, and sang a minor role.

FOREGROUND SECTION

It was nice not to be on stage all during the show.

In a way, there was more fun in only appearing briefly.

There was no need to take such a small character too seriously.

It meant less work rehearsing, and fewer worries.

BACKGROUND SECTION

The society put on a show like this two or three times a year.

Gilbert and Sullivan light comedies were the usual favourites.

The show would run for a week, with guaranteed audiences.

There was a lot of support for these performances.

TARGETS

She did not have a good enough voice for a big part.

She was very skilful with needle and thread.

1) Was it a professional opera company?

2) Where was the show staged?

(24) Several sixth-form boys were talking in the common room.

St. John's was a comprehensive, but still encouraged the sixth year.

Peter and Frank were describing a car they were building.

Mike was casually leafing through a magazine.

Terry was a senior prefect, and hoped to go to Cambridge.

FOREGROUND SECTION

A prefect's job is mainly concerned with discipline.

There were several unpleasant duties to be performed.

Staff cannot be everywhere, so the older pupils took charge.

The unruly younger children needed constant watching.

BACKGROUND SECTION

The school was a large, modern building in the suburbs.

Built before the education cuts, its facilities were excellent.

There were several large, well-tended sports fields.

The playgrounds were used as tennis courts in summer.

TARGETS

He often had to supervise the dinner queue.

He was wroking hard to get a scholarship.

- 1) Where were the boy's talking?
- 2) Was St. John's a grammar school?

APPENDIX 9

PRACTICE PASSAGES - EXPERIMENT 6

A. EXPLICIT PRACTICE PASSAGE

Up on the terminal roof, the view of the runways was marvellous. Telescopes were provided for the watchers, to look at the aircraft. Mrs. Maunders and her son Peter had been there for an hour, watching it all. He was fascinated by anything to do with aeroplanes. The only problem was the noise from the jets as they taxied just below. It was quieter in the observation lounge because of good double glazing.

- 1) Who was fascinated by anything to do with aeroplanes?
- 2) Why was it quieter in the observation lounge?

B. CONCEALED PRACTICE PASSAGES

- (1) On Saturday, Tim went for a stroll through the local street market. He liked the bustle and sense of life that filled the place. There was always something to see and hear, if you had the time. He felt sorry for harried shoppers, rushing to get their weekened provisions. They were unable to enjoy the market's essential charm and atmosphere. Tim stopped for a moment to listen to a salesman's stream of patter. The man claimed to be offering tea sets at a bargain price. He was very persuasive, trying to get people to buy the goods.

- 1) What were harried shoppers unable to enjoy in the market?
- 2) What was the salesman trying to get people to buy?

(2) Mrs. Warren prepared for her annual visit to her daughter with great care.

Despite going there each year, she never got used to the trip.

First she packed a large suitcase, with clothes and things for her stay.

Then she filled the smaller bag, with all the items for the train.

Magazines, sandwiches, flask, paper tissues, travel sickness pills.

She was nearly seventy, so perhaps her ritual precautions were excusable.

It was a long trip, and she had to change trains on two occasions.

Her worst nightmare was of being stranded by a missed connection.

1) How old was Mrs. Warren?

2) What was her worst nightmare?

APPENDIX 10

NEW EXPERIMENTAL PASSAGE FOREGROUNDING

BLOCKS - EXPERIMENT 7

Passage numbers given here refer to the passages given in full
in Appendix 8

(13) It was never necessary to get drunk in order to have a good time.

Soft drinks, with no ill effects, were more pleasant and less harmful.

In this way, life could be enjoyed to the full.

They ensured a clear head at night and no hangover next morning.

(5) The frozen slush at the side of the road was like a skating rink.

It was a real hazzard, especially for people in a hurry.

There was no way for shoes to get a grip on the surface.

High heels, of course, only made matters worse.

(7) The staff in restaurants come and go quite frequently.

The high staff turnover was inevitable in that kind of job.

There was no job security, and wages were often paid in cash.

This meant that people could move around quite freely.

(21) It was pleasant to have a smoke while relaxing.

Despite all the health warnings, it was a difficult habit to break.

In free moments, lighting up was an automatic reaction.

It seemed to add something to any short rest from work.

(9) A successful business depended on personal service and low prices.

The way to do well was to please the customers.

Income was guaranteed by having a large number of regulars.

Because of the good service, people kept on coming back.

(19) The soft Irish accent was blown away in the wind.

No-one was quite able to catch what was said.

To make matters worse, there was a lot of noise going on in the yard.

A low, mumbled voice was easily drowned out.

APPENDIX 11

PRACTICE AND DISTRACTOR PASSAGES -

EXPERIMENTS 7-9

A. EXPLICIT PRACTICE PASSAGE

On Saunday, Karen took her dog, Rex, for a walk in the park.

Rex was a large, boisterous Alsation, and needed plenty of exercise.

Walking through the streets to the park, Karen kept the dog on a tight lead.

Given half the chance, he would dash into the road, chasing motorbikes.

In the park, it was safe to let him run free for a while.

At first, Karen just walked along, letting Rex run around exploring.

Then she got out a rubber ball, and threw it for him to chase.

He would run after the ball all right, but often refused to bring it back.

When the time came to go home, she had to put the lead on quickly.

If he saw it coming, he would run away, and have to be caught.

1) What breed of dog was Rex?

2) What did Karen throw for him to chase?

B. CONCEALED PRACTICE PASSAGES

Suddenly there was a great roar of triumph from the North Road stand.

In the closing minutes of the match, the home team had at last scored a goal.

When the whistle blew at the end of the game, they were still in the lead.

The cheering and shouting from the local supporters was deafening.

As the football ground emptied, the police kept the rival fans apart.

Leo and Kevin made straight for the car park, behind the terraces.

Trevor had offered to drive them home, and he was waiting by his car.

He drove quickly out onto the road, anxious to avoid the rush of traffic. The three of them talked about the match, and about the club's prospects. Because they were feeling hungry, they pulled over at a fish and chip shop. They did not like the pies and hot dogs that were sold at the ground. Now, with the new alcohol rules, it was even impossible to get a drink there.

- 1) Was it the home team or the visiting team that won the match?
- 2) What food was sold at the football ground?

(2) Philip had applied for the job more or less as an act of desperation.

His paper qualifications were good, but he was unsure of his suitability.

Now, though, he had received a letter asking him to come for an interview.

This was encouraging, and he did his best to make a good impression.

He put on his suit, and wore a tie, for the first time in months.

The interviews were held at the company headquarters, a tall office block.

Half a dozen other candidates were waiting when he arrived.

One by one, a secretary called each applicant and led them out.

When Philip's turn came, he followed her nervously, tugging at his jacket.

There were three people sitting behind a table, with a chair in front of it.

Philip sat down, and they at once began asking him questions.

- 1) How many people were there on the interview panel?
- 2) What did Philip wear to make a good impression?

C. DISTRACTOR PASSAGES

(1) On the day of the local league snooker final, Frank Worth was taken ill. Harry, Don, Ian and the other members of his team were frantic. Sam Brayshaw, the league chairman, ruled they could pick a substitute. Only Jeff, a total novice, was available at such short notice. On the face of it, the team's prospects were not very hopeful. Nevertheless, both sides were determined to play a good game. He might after all manage to get some lucky shots. The first frame of the match began as the clock struck eight. By the time of the first interval, the scores were level. The conversation around the bar was lively and excited. Everyone was getting caught up in the finely balanced contest.

- 1) Who ruled that a substitute could be picked?
- 2) Where did the conversation take place during the interval?

(2) The streets of the small Spanish town thronged with people. Normally a sleepy village, at fiesta time it bustled with life. A group of English tourists gathered at a small pavement cafe. James and his wife Valerie were bullfight fans, and came every year. Albert was also a regular visitor to the festivities. Charles was there for the first time, and was unsure what to expect. He hoped the killing would not be too unpleasant to watch. Small children were letting off firecrackers in the streets. Old people winced unhappily at the loud, unexpected bangs.

- 1) Which of the tourists was at the fiesta for the first time?
- 2) Why were old people wincing unhappily?

(3) There was a slight mist that Sunday morning, but the sun was getting through.

It would be a grand day for the planned fishing trip.

Adam drove his car, loaded with their gear, right up to the river bank.

Gus and Nigel walked a little upstream, while Brian went the other way.

Ray was the only one without a fold-up seat, and had to sit on the ground.

As the early morning sun rose higher, steam curled up from the water.

There were small splashes as fish rose occasionally to the surface.

He was just a little uncomfortable, because of the wet dew on the grass.

At about noon, they all went to a nearby pub for lunch.

They ate bread, cheese and pickles, with a pint of beer to wash it down.

1) Whose car did the fishermen drive to the river in?

2) Where did they go for lunch?

(4) In the autumn, Graham liked to take arly morning walks along the cliffs.

During the summer, there were too many tourists for it to be much fun.

He would start out just as it was getting light, and make for North Head.

This was a large headland just outside the town.

The path wound its way around the edge of a sheer drop to the sea below.

It was dangerous if there was a strong wind, as you could be blown over.

From the headland, Graham would follow the path down onto the beach.

He would usually look for driftwood, or anything left by the tide.

One day, he found the beach littered with seabirds covered in oil.

There were patches of oil on the sand, and more still floating in.

He ran at once to call the police and report his discovery.

This had happened once before, and hundreds of birds had died.

1) Why was it not much fun to walk along the cliffs in summer?

2) What did Graham usually look for on the beach.

APPENDIX 12

EXPERIMENTAL PASSAGES - EXPERIMENTS 8 - 9

Each passage is followed by the two target sentences used in Experiment 9, with slash marks to indicate the division into initial and final segments.

(1) On Sunday morning, Rayner Road was busy with people doing odd jobs.

Everyone was proud that the houses were all so well cared for.

The street had once won a prize, awarded by the local newspaper.

Such a well-tended appearance was a credit to the town.

At the end of the street, Peter Tyrell was painting the front door.

Further down, Steve at number 15 was sweeping the driveway.

Next door, John was gardening, while further on Jeff was cleaning the windows.

· FOREGROUND 1

The garden was a big one, with many flower beds and small bushes.

With spring rapidly advancing, the lawn would have to be mowed.

Besides this, the rockery looked as though it could do with some attention.

There was a large vegetable patch to be turned over, ready for planting.

FOREGROUND 2

The panes of glass were thoroughly dirty, and needed a good wash.

Warm soapy water was able to cut through the grime with ease.

Energetic rubbing with a dry cloth then began to bring back the shine.

The glass started to sparkle in the morning sun.

BACKGROUND

It was a bright, sunny spring morning, with a warm, gentle breeze.

The birds were singing, and everywhere seemed very peaceful.

Church bells could faintly be heard, coming from far in the distance.

Only a few fluffy white clouds were dotted in the deep blue sky.

TARGETS

He went to fetch a large digging fork from the shed.

He climbed up a long ladder to reach to the first floor.

1) Why had the street once won a prize?

2) Who had awarded the prize?

He went to fetch a // large digging fork from the shed.

He used a long // ladder to reach to the first floor.

(2) On Friday night, after school, Mick went down to the fairground.

Brendan and Sheila were waiting at the gates, and they went in together.

The fair was in one corner of the huge municipal park.

There were tennis courts and a mini-golf course in the park as well.

In a few minutes, some more of their friends arrived.

For a while, they all went round in a group, but then they split up.

Glenda and Julie wanted to go on the dodgems, and so did Martin.

Clare went to the hall of mirrors and June had a ride on the big dipper.

FOREGROUND 1

It was a great joke to see yourself reflected in ridiculous, ugly ways.

Sometimes the reflection was fat, and sometimes as thin as a rake.

By moving slightly, the features could be distorted even more.

There were a lot of mirrors to look into, all of different types.

FOREGROUND 2

The posters claimed that this was the biggest roller-coaster in the country.

It was long and twisting, and in places climbed very high indeed. From the highest point of the ride, it was possible to see for miles. The speeds attained were terrific, with the wind whistling past.

BACKGROUND

It was early in the evening, and the fair was still relatively quiet. The glowing lights looked dim and faint in the strong twilight. Later, the crowds would build up and the place would become more lively.

The fair was one of the most popular summer events in that town.

TARGETS

She giggled at some of the funniest of the reflections.

She shrieked in terror at every sudden, unexpected drop.

- 1) What day of the week was it?
- 2) Why did the group eventually split up?

She could not stop // laughing at the funny reflections.

She looked forward most // to the sudden unexpected drops.

- (3) The local amateur operatic society was preparing for another show. They put on at least two shows a year, and they were very popular. The company had been going for fifteen years now, with much success. All the shows were a joint effort, with everyone doing their bit. Keith and Norman would paint the scenery and get the stage ready. Mrs. Hindle, always did the refreshments, while Jane handled publicity. Hilda sewed all the costumes, and Christine sang the main female leads.

FOREGROUND 1

All the clothes for the shows had to be stitched up from old material. It took real ingenuity to piece something together out of waste cloth. Careful stitching was essential, especially when doing the gowns. Besides their appearance, the clothes also had to be made to fit.

FOREGROUND 2

Playing such important roles was always a tiring experience. Being on stage so long made each production seem like hard work. The songs all had to be learned perfectly, and there was a lot to rehearse. Still, a beautiful voice was a gift which had to be used.

BACKGROUND

The shows were always staged in the local civic hall. The facilities there were excellent, for performances of all kinds. There was ample room for audiences, with many tiers of upholstered

seats.

Good spotlights were fixed all round the walls, controlled automatically.

TARGETS

She was very proud of creating good quality outfits.

She always received loud cheers for excellent singing.

1) How many shows a year did the company perform?

2) How long had the company been in existence?

She was very proud // of creating good quality outfits.

She always received // loud cheers for excellent singing.

(4) On Bank Holiday Monday, the Carters set off on a day trip.

Mr. Carter was off work, so he could drive the family to the seaside.

They started early, packing the car with picnic things.

At first the driving was easy, but nearer the coast cars jammed the roads.

When they reached their destination, the Carters went down to the beach.

Mrs. Carter took Matthew, the youngest child, for a donkey ride.

Simon went off to the amusement arcade.

Adam made some sand-castles while Gregory went for a swim.

FOREGROUND 1

The sand was soft and moist, just right for building.

The castle had to have a moat, of course, dug out with a spade.

Lollipop sticks could be used as flagpoles on the battlements.
Pebbles pressed into the sides added an extra decoration.

FOREGROUND 2

The water was cool for the time of the year, but pleasantly so.
As the waves came tumbling in, it was fun to dive under the foam.
The water offshore was shallow, and it was safe to go out a long way.
It would only be dangerous if the tide was on the ebb.

BACKGROUND

It was a hot day, with hardly a breath of wind to stir the air.
The sun was scorching down, creating a fierce glare.
Cold beers and ice creams were selling well.
It was almost too hot to sunbathe in comfort.

TARGETS

He filled a bucket to make another turret.
He splashed and dived through the breakers.

- 1) Why was Mr. Carter off work that day?
- 2) What did they pack the car with?

He really // enjoyed designing turrets and towers.
He spent all afternoon // diving through the breakers.

(5) On Saturday night, a group of friends met in the pub.

They had virtually grown up together, living in one neighbourhood.

Although still friends, they had all gone their separate ways.

Gary had joined the army, and was home on leave.

Theresa was a student, and Clive worked as a bank clerk.

Carol was a qualified nurse, and Pat was a typist in an insurance firm.

FOREGROUND 1

Working in the health service was a demanding profession.

Caring for the sick was hard work, and emotionally draining.

The hours were long, and the patients needed constant attention.

It took total dedication to be able to stand the pace.

FOREGROUND 2

It was very boring to spend all day typing.

Some of the documents were complex, with long and technical words.

Sitting at the typewriter seemed to put a strain on the back.

Trying to copy from poor handwriting made matters even worse.

BACKGROUND

It was a nice little pub, with a very pleasant atmosphere.

In the corner, a small open fire was burning.

The ceiling was criss-crossed with dark oak beams.

The tables and chairs were also of old, sturdy wood.

TARGETS

She was working on a long-stay surgical ward.

She especially hated doing long business letters.

1) Why were the group of people such good friends?

2) What day of the week was it?

She was working on // a long-stay surgical ward.

She found it difficult // to do long business letters.

(6) The Garners lived in a large detached house outside the town.

Both parents worked, and they employed a housekeeper and a handyman.

At weekends, with everyone home, the house was a hive of activity.

Of the five children, Tim was keen on sports.

Joanne's hobby was photography, and Marl made model aircraft.

Paul loved playing the violin, while Frank watched television all day long.

FOREGROUND 1

Music was a fine interest, and gave much pleasure.

The strings are among the most moving of the instruments.

It was pleasant to work hard on some difficult piece, and get it just right.

By determined effort, real skill at playing could be acquired.

FOREGROUND 2

Lolling in front of a screen all day is almost criminal.

Even in fine weather, the programmes seemed too fascinating to leave.

In reality, of course, it was just plain laziness.

Television is a drug, and it was easy to become totally addicted.

BACKGROUND

The house was on a broad, tree-lined avenue, in a very respectable area.

There were no busy main roads nearby, and it was a quiet place.

Because it was such a good area, house prices were high.

Three-car garages were not an unusual sight.

TARGETS

He was hoping to join the regional youth orchestra.

He would watch anything from serials to documentaries.

1) What was the name of the family?

2) What domestic staff did they employ?

He was hoping to // join the regional youth orchestra.

He was most interested in // serials and documentaries.

APPENDIX 13

EXAMPLE SEGMENTED PASSAGE - EXPERIMENT 9

Slash marks are used to indicate division into sentence segments.

This is passage (2), from Appendix 12.

On Friday night // after school Mick went // down to the fairground.
Brendan and Sheila // were waiting // at the gates and // they
went in together.

The fair was // in one corner of // the huge municipal park.
There were tennis // courts and // a mini-golf course in // the
park as well.

In a few minutes // some more of their friends arrived.
For a while they // all went round in a // group but then they
split up.

Glenda and Julie wanted to go // on the dodgems and // so did Martin.
Clare went to the // hall of mirrors and // June had a ride //
on the big dipper.

FOREGROUND 1

It was a great // joke to // see yourself reflected // in ridiculous
ugly ways.

Sometimes the reflection was fat and sometimes // as thin'// as a rake.
By moving // slightly the features could // be distorted even more.
There were a lot // of mirrors to look // into all // of different types.

FOREGROUND 2

The posters claimed that // this was the biggest roller-coaster in
the country.

It was // long and twisting and in places // climbed very high indeed.
From the // highest point of // the ride it was // possible to
see for miles.

The speeds attained were terrific with the wind // whistling past.

BACKGROUND

It was early // in the // evening and the fair // was still relatively quiet.

The slowing lights looked // dim and faint in the strong twilight.

Later the crowds would // build up and the // place would // become more lively.

The fair was // one of the // most popular summer // events in that town.

TARGETS

She could not stop // laughing at the funny reflections.

She looked forward most // to the sudden unexpected drops.

- 1) What day of the week was it?
- 2) Why did the group eventually split up?

APPENDIX 14

ANOVA SUMMARY TABLES - EXPERIMENT 1

1) READER SKILL X DISTANCE X REFERENCE TYPE

	SOURCE OF VARIATION	DF	SS	SSX	MS	VR
32	RDR.SUB STRATUM					
32	RDR	1	4.310E	2.90	4.310E	1.712
33	RESIDUAL	22	5.539E	37.21	2.518E	11.790
34	TOTAL	23	5.970E	40.11	2.596E	12.155
38	RDR.SUB.PRO STRATUM					
38	PRO	1	6.233E	0.00	6.233E	0.032
39	RDR.PRO	1	2.930E	0.02	2.930E	0.152
40	RESIDUAL	22	4.240E	2.85	1.927E	0.903
40	TOTAL	24	4.276E	2.87	1.782E	0.834
42	RDR.SUB.DIS STRATUM					
42	DIS	2	3.943E	0.03	1.971E	0.072
44	RDR.DIS	2	5.904E	0.40	2.952E	1.074
45	RESIDUAL	44	1.207E	8.12	2.748E	1.287
46	TOTAL	48	1.272E	8.55	2.650E	1.241
49	RDR.SUB.PRO.DIS STRATUM					
49	PRO.DIS	2	1.112E	0.75	5.562E	2.608
50	RDR.PRO.DIS	2	1.598E	0.11	7.981E	0.374
51	RESIDUAL	44	9.385E	6.30	2.133E	0.999
52	TOTAL	48	1.866E	7.16	2.220E	1.040
54	RDR.SUB.PRO.DIS. UNITS STRATUM					
54		288	6.150E	41.32	2.135E	6
56	GRAND TOTAL	431	1.489E	100.00		

2) READER SKILL X DISTANCE - PRONOUN SENTENCES ONLY

SOURCE OF VARIATION	DF	SS	SSX	PS	VR
RDR•SUB STRATUM					
RDR	1	6048243	4.31	6648243	1.318
RESIDUAL	22	100935712	71.95	4587586	6.796
TOTAL	23	106983952	76.26	4651476	6.850
RDR•SUB•DIS STRATUM					
DIS	2	1339960	0.96	669980	0.992
RER•DIS	2	2260893	1.61	1130446	1.675
RESIDUAL	44	29704056	21.17	675092	
TOTAL	48	33304908	23.74	693852	
GRAND TOTAL	71	140288832	100.00		

3) SKILL X SERIAL POSITION

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
RDR•SUB STRATUM					
RDR	1	41086792	33.63	41086792	25.912
RESIDUAL	22	34882792	28.56	1585581	6.756
TOTAL	23	75969584	62.19	3303025	14.074
RDR•SUB•POS STRATUM					
PCS	5	13571188	11.11	2714238	11.565
RDR•POS	5	6800698	5.57	1360029	5.795
RESIDUAL	110	25816748	21.13	234698	
TOTAL	120	46188032	37.81	384900	
GRAND TOTAL	143	122157616	100.00		

APPENDIX 15

ANOVA SUMMARY TABLES - EXPERIMENT 2

1) SKILL X POSITION X DISTANCE - PRONOUN READING TIMES

SOURCE OF VARIATION	DF	SS	SS%	MS	VF
RDR•SUB STRATUM					
RDR	1	75319984	17.20	75319984	12.802
RESIDUAL	22	129431104	29.55	5883222	9.156
TOTAL	23	204751072	46.74	8912220	12.914
RDR•SUB•POS STRATUM					
PCS	1	9953594	2.27	9953594	11.463
RDR•POS	1	15472	0.00	15472	0.018
RESIDUAL	22	19122648	4.36	868702	1.357
TOTAL	24	29671712	6.64	1211321	1.853
RDR•SUB•DIS STRATUM					
DIS	4	39433894	6.95	7608451	7.512
RDR•DIS	4	4433915	1.01	1108479	1.054
RESIDUAL	88	89122752	20.35	1012872	1.583
TOTAL	96	124000464	28.31	1291672	2.015
RDR•SUB•POS•DIS STRATUM					
PCS•DIS	4	15652972	3.57	3913468	6.117
RDR•POS•DIS	4	8245714	1.88	2061429	3.222
RESIDUAL	88	56311248	12.85	639787	
TOTAL	96	80200832	18.31	835425	
GRAND TOTAL	239	428024364	100.00		

2) SKILL X POSITION X DISTANCE - PRONOUN QUESTION ACCURACY

	SOURCE OF VARIATION	DF	SS	SSX	MS	VR
28						
30	RNR.SUR STRATUM					
	RNR	1	0.0667	0.07	0.0667	0.118
32	RESIDUAL	22	12.4667	12.60	0.5667	1.678
	TOTAL	23	12.5333	12.67	0.5449	1.614
34	RDR.SUB.POS STRATUM					
	POS	1	1.3500	1.36	1.3500	4.223
36	RDR.POS	1	0.8167	0.83	0.8167	2.555
38	RESIDUAL	22	7.0333	7.11	0.3197	0.947
	TOTAL	24	9.2000	9.30	0.3833	1.135
40	RDR.SUB.DIS STRATUM					
	DIS	4	10.1417	10.25	2.5354	7.125
42	RDR.DIS	4	0.8083	0.82	0.2021	0.727
44	RESIDUAL	88	24.4500	24.71	0.2778	0.823
	TOTAL	96	35.4000	35.78	0.3687	1.092
46	RDR.SUB.POS.DIS STRATUM					
	POS.DIS	4	10.7750	10.89	2.6937	7.977
48	RDR.POS.DIS	4	1.3083	1.32	0.3271	0.969
50	RESIDUAL	88	29.7167	30.04	0.3377	
	TOTAL	96	41.8000	42.25	0.4354	
52	GRAND TOTAL	239	98.9333	100.00		

3) SKILL X QUESTION TYPE: - SECOND QUESTION ACCURACY

	SOURCE OF VARIATION	DF	SS	SSX	MS	VR
28						
30	RDR•SUB STRATUM					
	RDR	1	0.188	0.25	0.188	0.102
32	RESIDUAL	22	40.292	53.03	1.831	1.562
	TOTAL	23	40.479	53.28	1.760	1.501
34	RDR•SUB•INFER STRATUM					
	INFER	1	9.188	12.09	9.188	7.837
36	RDR•INFER	1	0.521	0.69	0.521	0.444
	RESIDUAL	22	25.792	33.95	1.172	
38	TOTAL	24	35.500	46.72	1.479	
40	GRAND TOTAL	47	75.979	100.00		

4) SKILL X TEST TYPE - RECOGNITION TEST LATENCIES

SOURCE OF VARIATION	LF	SS	SS%	FS	VF
RDR•SUB STRATUM					
PER	1	53485712	12.11	92485712	5.648
RESIDUAL	22	384671936	47.26	16875596	2.371
TOTAL	23	458157632	59.37	10919096	2.848
RDR•SUB•CONE STRATUM					
CCNC	2	2535512	0.33	1269756	0.182
RDR•COND	2	3204383	0.42	1602192	0.223
RESIDUAL	44	387753624	39.88	6894386	
TOTAL	48	313496896	40.63	6521185	
GRAND TOTAL	71	771654528	100.00		

5) SKILL X TEST TYPE - RECOGNITION TEST ACCURACY

	SOURCE OF VARIATION	DF	SS	SSX	MS	VR
28	RDR.SUB STRATUM					
30	RDR	1	0.0139	0.04	0.0139	0.054
	RESIDUAL	22	5.6389	16.43	0.2563	0.883
32	TOTAL	23	5.6528	16.47	0.2458	0.846
34	RDR.SUB.NAT STRATUM					
	NAT	2	15.0278	43.79	7.5139	25.874
36	RDR.NAT	2	0.8611	2.51	0.4306	1.483
	RESIDUAL	44	12.7778	37.23	0.2904	
38	TOTAL	48	28.6667	83.53	0.5972	
40	GRAND TOTAL	71	34.3194	100.00		

APPENDIX 16

ANOVA SUMMARY TABLES - EXPERIMENT 3

1) SKILL X DISTANCE X FOREGROUNDING - PRONOUN SENTENCE READING TIMES

SOURCE OF VARIATION	DF	SS	SS%	MS	VF
RDR•SUB STRATUM					
RDR	1	95363888	9.07	85363888	7.558
RESIDUAL	22	248465760	26.39	11293898	4.264
TOTAL	23	333829632	35.46	14514330	5.480
RDR•SUB•FOR STRATUM					
FOR	1	54241216	5.76	54241216	31.401
RDR•FOR	1	222711	0.02	222711	0.129
RESIDUAL	22	38002368	4.04	1727380	0.652
TOTAL	24	92466288	9.82	3852762	1.455
RDR•SUB•DIS STRATUM					
DIS	4	55026472	5.85	13756618	6.242
RDR•DIS	4	2834342	0.30	708586	0.322
RESIDUAL	88	193930848	20.60	2203760	0.832
TOTAL	96	251791648	26.75	2622830	0.998
RDR•SUB•FOR•DIS STRATUM					
FOR•DIS	4	26277876	2.79	6569469	2.480
RDR•FOR•DIS	4	3910147	0.42	977537	0.369
RESIDUAL	88	233094688	24.76	2648803	
TOTAL	96	263282688	27.97	2742528	
GRAND TOTAL	239	941370112	100.00		

2) SKILL X CONDITIONS - PRONOUN SENTENCE READING TIMES

SOURCE OF VARIATION	DF	SS	SS%	MS	VF
RDR.SUR STRATUM					
ERR	1	82017296	9.27	82017296	7.496
RESIDUAL	22	249714016	27.21	10941546	4.397
TOTAL	23	322731264	36.48	14031794	5.638
RDR.SUR.COND STRATUM					
COND	8	117948912	13.23	14743614	5.924
RDR.COND	3	6098667	0.69	762223	0.306
RESIDUAL	176	437992064	49.50	2488591	
TOTAL	192	562030552	63.52	2927289	
GRAND TOTAL	215	884770816	100.00		

3) SKILL X FOREGROUNDING X DISTANCE - PRONOUN QUESTION ACCURACY

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
RDR SUB STRATUM					
RDR	1	0.1042	0.14	0.1042	0.186
RESIDUAL	22	12.2917	16.07	0.5587	1.638
TOTAL	23	12.3958	16.20	0.5389	1.580
RDR SUB FOR STRATUM					
FOR	1	0.0375	0.05	0.0375	0.103
RDR FOR	1	0.8375	0.05	0.8375	0.103
RESIDUAL	22	8.0250	10.49	0.3648	1.069
TOTAL	24	8.8800	10.59	0.3375	0.985
RDR SUB DIS STRATUM					
DIS	4	2.7250	3.56	0.6812	2.843
RDR DIS	4	0.7917	1.03	0.1979	0.826
RESIDUAL	88	21.0833	27.56	0.2396	0.702
TOTAL	96	24.6000	32.16	0.2553	0.751
RDR SUB FOR DIS STRATUM					
FOR DIS	4	0.3583	0.47	0.0896	0.263
RDR FOR DIS	4	1.0250	1.34	0.2562	0.751
RESIDUAL	88	30.0167	39.24	0.3411	
TOTAL	96	31.4000	41.05	0.3271	
GRAND TOTAL	239	16.4958	100.00		

APPENDIX 17

ANOVA SUMMARY TABLES - EXPERIMENT 4

1) SKILL X FOREGROUNDING X DISTANCE X TRIAL - RECOGNITION TEST LATENCIES

28	SOURCE OF VARIATION	DF (MV)	SS	SSX	MS	VR
30	RDR.SUB STRATUM					
	RDR	1	5.626E	5	0.03	5.626E 5 0.047
32	RESIDUAL	22	2.623E	8	14.35	1.192E 7 2.619
34	TOTAL	23	2.628E	8	14.38	1.143E 7 2.510
36	RDR.SUB.FOR STRATUM					
	FOR	1	4.168E	6	0.23	4.168E 6 2.686
	RDR.FOR	1	1.420E	6	0.08	1.420E 6 0.915
38	RESIDUAL	22	3.414E	7	1.87	1.552E 6 0.341
40	TOTAL	24	3.973E	7	2.17	1.655E 6 0.364
42	RDR.SUB.DIS STRATUM					
	DIS	4	2.013E	7	1.10	5.032E 6 1.077
	RDR.DIS	4	1.479E	7	0.81	3.697E 6 0.791
44	RESIDUAL	88	4.112E	8	22.50	4.673E 6 1.026
46	TOTAL	96	4.461E	8	24.41	4.647E 6 1.021
48	RDR.SUB.TRIAL STRATUM					
	TRIAL	1	1.705E	7	0.93	1.705E 7 2.243
	RDR.TRIAL	1	2.534E	7	1.39	2.534E 7 3.335
50	RESIDUAL	22	1.672E	8	9.15	7.598E 6 1.669
52	TOTAL	24	2.095E	8	11.47	8.731E 6 1.918
54	RDR.SUB.FOR.DIS STRATUM					
	FOR.DIS	4	1.814E	7	0.99	4.536E 6 1.031
	RDR.FOR.DIS	4	1.798E	7	0.98	4.495E 6 1.021
56	RESIDUAL	87(1)	3.829E	8	20.95	4.401E 6 0.967
	TOTAL	95	4.190E	8	22.93	4.411E 6 0.969

RDR.SUB.FOR.TRIAL STRATUM									
60	FOR.TRIAL	1	6.061E	6	0.33	6.061E	6	0.767	
	RDR.FOR.TRIAL	1	1.590E	7	0.87	1.590E	7	2.012	
62	RESIDUAL	22	1.738E	8	9.51	7.899E	6	1.735	
64									
2									
4									
6	TOTAL	24	1.957E	8	10.71	8.155E	6	1.791	
8	RDR.SUB.DIS.TRIAL STRATUM								
	DIS.TRIAL	4	1.667E	7	0.91	4.166E	6	1.440	
10	RDR.DIS.TRIAL	4	5.358E	7	2.93	1.340E	7	4.631	
	RESIDUAL	88	2.546E	8	13.93	2.893E	6	0.635	
12	TOTAL	96	3.248E	8	17.77	3.383E	6	0.743	
14	RDR.SUB.FOR.DIS.TRIAL STRATUM								
	FOR.DIS.TRIAL	4	2.211E	7	1.21	5.528E	6	1.214	
16	RDR.FOR.DIS.TRIAL	4	6.862E	6	0.38	1.716E	6	0.377	
	RESIDUAL	65(23)	2.959E	8	16.19	4.553E	6		
18	TOTAL	73	3.249E	8	17.78	4.451E	6		
20	GRAND TOTAL								
		455	2.223E	9	121.63				

2) SKILL X CONDITIONS - RECOGNITION TEST LATENCIES

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
RDR•SUB STRATUM					
RDR	1	224685	0.04	224685	0.023
RESIDUAL	22	213936576	35.53	9724388	4.951
TOTAL	23	214161248	35.57	9311358	4.741
RDR•SUB•COND STRATUM					
COND	8	27475120	4.56	3434390	1.749
RDR•COND	8	14827786	2.46	1853473	0.944
RESIDUAL	176	345654592	57.41	1963947	
TOTAL	192	387957440	64.43	2020612	
GRAND TOTAL	215	602118656	100.00		

3) SKILL X FOREGROUNDING X DISTANCE - RECOGNITION TEST ACCURACY

	SOURCE OF VARIATION	DF	SS	SS%	MS	VR
28	RDR.SUB STRATUM					
30	RDR	1	0.00000	0.00	0.00000	0.000
32	RESIDUAL	22	1.80000	7.63	0.08182	0.896
	TOTAL	23	1.80000	7.63	0.07826	0.857
34	RDR.SUB.FOR STRATUM					
36	FOR	1	0.15000	0.64	0.15000	2.020
38	RDR.FOR	1	0.41667	1.77	0.41667	5.613
	RESIDUAL	22	1.63333	6.92	0.07424	0.813
	TOTAL	24	2.20000	9.32	0.09167	1.884
40	RDR.SUB.DIS STRATUM					
42	DIS	4	0.68333	2.90	0.17083	1.550
	RDR.DIS	4	0.41667	1.77	0.10417	0.945
	RESIDUAL	88	9.70000	41.10	0.11023	1.207
	TOTAL	96	10.80000	45.76	0.11250	1.232
46	RDR.SUB.FOR.DIS STRATUM					
48	FOR.DIS	4	0.43333	1.84	0.10833	1.187
	RDR.FOR.DIS	4	0.33333	1.41	0.08333	0.913
	RESIDUAL	88	8.03333	34.04	0.09129	
	TOTAL	96	8.80000	37.29	0.09167	
52	GRAND TOTAL	239	23.59999	100.00		

APPENDIX 18

ANOVA SUMMARY TABLES - EXPERIMENT 5

1) SKILL X FOREGROUNDING X DISTANCE - PRONOUN SENTENCE READING TIMES

	SOURCE OF VARIATION	DF	SS	SS%	MS	VR
28						
30						
	RDR.SUB STRATUM					
32	RDR	1	3255700	1.83	3255700	0.761
	RESIDUAL	22	94119648	52.95	4278165	6.463
34	TOTAL	23	97375344	54.78	4233710	6.395
36	RDR.SUB.FOR STRATUM					
	FOR	1	22768216	12.81	22768216	21.515
38	RDR.FOR	1	1027548	0.58	1027548	0.971
	RESIDUAL	22	23281668	13.10	1058258	1.590
40	TOTAL	24	47077432	26.48	1961560	2.963
42	RDR.SUB.DIS STRATUM					
	DIS	1	381276	0.21	381276	0.486
44	RDR.DIS	1	183575	0.10	183575	0.234
	RESIDUAL	22	17264048	9.71	784729	1.185
46	TOTAL	24	17820896	10.03	742871	1.122
48	RDR.SUB.FOR.DIS STRATUM					
	FOR.DIS	1	899581	0.51	899581	1.359
50	RDR.FOR.DIS	1	13944	0.01	13944	0.021
	RESIDUAL	22	14563808	8.19	661991	
52	TOTAL	24	15477332	8.71	644889	
54	GRAND TOTAL	95	177758976	100.00		

2) SKILL X FOREGROUNDING X DISTANCE - PRONOUN QUESTION ACCURACY

	SOURCE OF VARIATION	DF	SS	SS%	MS	VR
28						
30						
32	RDR.SUB STRATUM					
	RDR	1	0.6104	0.04	0.6104	0.04
	RESIDUAL	22	5.3958	18.51	0.2453	0.94
34	TOTAL	23	5.4062	18.54	0.2351	0.90
36	RDR.SUB.FOR STRATUM					
	FOR	1	0.5104	1.75	0.5104	1.28
38	RDR.FOR	1	0.6104	0.04	0.6104	0.02
	RESIDUAL	22	8.7292	29.74	0.3968	1.52
40	TOTAL	24	9.2500	31.73	0.3854	1.48
42	RDR.SUB.DIS STRATUM					
	DIS	1	0.0104	0.04	0.0104	0.02
44	RDR.DIS	1	0.6104	0.04	0.6104	0.02
	RESIDUAL	22	8.2292	28.22	0.3741	1.43
46	TOTAL	24	8.2500	28.30	0.3437	1.32
48	RDR.SUB.FOR.DIS STRATUM					
	FOR.DIS	1	0.2604	0.89	0.2604	1.00
50	RDR.FOR.DIS	1	0.2604	0.89	0.2604	1.00
	RESIDUAL	22	5.7292	19.65	0.2604	
52	TOTAL	24	6.2500	21.44	0.2604	
54	GRAND TOTAL	95	29.1562	100.00		

3) SKILL X FOREGROUNDING X DISTANCE X TRIAL - RECOGNITION TEST LATENCIES

SOURCE OF VARIATION	DF (MV)	SS	SS%	MS	VR
RDR.SUP STRATUM					
FOR	1	10522216	2.47	10522216	1.901
RESIDUAL	22	121743824	28.56	5533810	3.053
TOTAL	23	132266032	31.03	5750697	3.173
RDR.SUP.FOR STRATUM					
FOR	1	28071170	0.66	28071170	1.275
RDR.FOR	1	2696408	0.63	2696408	1.225
RESIDUAL	22	48439352	11.36	2201789	1.215
TOTAL	24	53942928	12.65	2247622	1.240
RDR.SUP.DIS STRATUM					
DIS	1	11647014	2.73	11647014	3.288
RDR.DIS	1	806467	0.19	806467	0.225
RESIDUAL	22	77941680	18.28	3542804	1.955
TOTAL	24	90395152	21.20	3766465	2.078
RDR.SUP.TRIAL STRATUM					
TRIAL	1	6131549	1.44	6131549	6.242
RDR.TRIAL	1	182313	0.04	182313	0.186
RESIDUAL	22	21609268	5.07	982239	0.542
TOTAL	24	27923128	6.55	1163464	0.642
RDR.SUP.FOR.DIS STRATUM					
FOR.DIS	1	4200360	0.99	4200360	1.465
RDR.FOR.DIS	1	439245	0.10	439245	0.153
RESIDUAL	21(1)	60229904	14.13	2868091	1.582
TOTAL	23	64869504	15.22	2820413	1.556

3) - CONTINUED

RDK.SUB.FOR.TRIAL STRATUM				
FOR.TRIAL	1	1616221	0.38	1616221
RDP.FOR.TRIAL	1	3062	0.00	3062
RESIDUAL	22	27443816	6.44	1247446
TOTAL	24	29063096	6.82	1210962
				1.296
				0.002
				0.688
				7.668

RDR.SUE.DIS.TRIAL STRATUM				
DIS.TRIAL	1	578052	0.14	578052
RDR.DIS.TRIAL	1	147167	0.03	147167
RESIDUAL	21(1)	20403888	4.79	971614
TOTAL	23	21129104	4.96	918657
				0.595
				0.151
				0.536
				0.507

RDP.SUB.FOR.DIS.TRIAL STRATUM				
FOR.DIS.TRIAL	1	3597537	0.84	3597537
RDR.FOR.DIS.TRIAL	1	711469	0.17	711469
RESIDUAL	14(R)	25373700	5.95	1812407
TOTAL	16	29682734	6.96	1855169
				1.985
				0.393

GRAND TOTAL	181	449271488	105.39	
-------------	-----	-----------	--------	--

4) SKILL X FOREGROUNDING X DISTANCE - RECOGNITION TEST ACCURACY

	SOURCE OF VARIATION	DF	SS	SS%	MS	VR
28						
30						
32	RDR.SUB STRATUM					
	RDR	1	0.0417	0.38	0.0417	0.208
	RESIDUAL	22	4.4167	40.30	0.2008	1.828
34	TOTAL	23	4.4583	40.68	0.1938	1.765
36	RDR.SUB.FOR STRATUM					
	FOR	1	0.0417	0.38	0.0417	0.647
38	RDR.FOR	1	0.0417	0.38	0.0417	0.647
	RESIDUAL	22	1.4167	12.93	0.0644	0.586
40	TOTAL	24	1.5000	13.69	0.0625	0.569
42	RDR.SUB.DIS STRATUM					
	DIS	1	0.0417	0.38	0.0417	0.379
44	RDR.DIS	1	0.0417	0.38	0.0417	0.379
	RESIDUAL	22	2.4167	22.05	0.1098	1.000
46	TOTAL	24	2.5000	22.61	0.1042	0.946
48	RDR.SUB.FOR.DIS STRATUM					
	FOR.DIS	1	0.0417	0.38	0.0417	0.379
50	RDR.FOR.DIS	1	0.0417	0.38	0.0417	0.379
	RESIDUAL	22	2.4167	22.05	0.1098	
52	TOTAL	24	2.5000	22.61	0.1042	
54	GRAND TOTAL	95	10.9583	100.00		

APPENDIX 19

ANOVA SUMMARY TABLES - EXPERIMENT 6

1) SKILL X FOREGROUNDING X RELATEDNESS X DISTANCE - PRONOUN SENTENCE

READING TIMES

SOURCE OF VARIATION	DF	SS	SSX	MS	VR
RDR•SUB STRATUM					
RDR	1	34309904	6.53	34309904	4.360
RESIDUAL	22	173126528	32.96	7869207	10.942
TOTAL	23	207436416	39.50	9018974	12.541
RDR•SUB•DIS STRATUM					
DIS	2	12043816	2.29	6021908	4.407
RDR•DIS	2	2449251	0.47	1224626	0.89E
RESIDUAL	44	60122336	11.45	1366417	1.900
TOTAL	48	74615392	14.21	1554427	2.162
RDR•SUB•FCR STRATUM					
FOR	1	13334090	2.54	13334090	7.586
RDR•FOR	1	381246	0.07	381246	0.217
RESIDUAL	22	38670888	7.36	1757768	2.444
TOTAL	24	52386224	9.97	2182759	3.035
RDR•SUB•REL STRATUM					
REL	1	994697	0.19	994697	0.751
RDR•REL	1	471704	0.09	471704	0.35E
RESIDUAL	22	29144456	5.55	1324748	1.842
TOTAL	24	30610856	5.83	1275452	1.774
RDR•SUB•DIS•FCR STRATUM					
DIS•FOR	2	4617366	0.88	2308683	2.129
RDR•DIS•FCR	2	1014352	0.19	507196	0.468
RESIDUAL	44	47711192	9.08	1084345	1.508
TOTAL	48	53342944	10.16	1111311	1.545

✓ RDR.SUB.DIS.REL STRATUM
DIS.REL
RDR.DIS.REL
RESIDUAL
TOTAL

2	395060	0.08	197230	0.165
2	2641171	0.50	1320586	1.101
44	52778016	10.05	1199500	1.662
48	55814240	10.63	1122757	1.617

RDR.SUB.FOR.REL STRATUM
FCR.REL
RDR.FOR.REL
RESIDUAL
TOTAL

1	5559303	1.06	5559303	9.913
1	29131	0.01	29131	0.052
22	12338372	2.35	560835	0.780
24	17926804	3.41	746550	1.035

RDR.SUB.DIS.FCR.REL STRATUM
DIS.FOR.REL
RDR.DIS.FCR.REL
RESIDUAL
TOTAL

2	1280916	0.24	640458	0.891
2	143246	0.03	71623	0.100
44	31643464	6.03	715170	
48	33067624	6.30	688909	

GRAND TOTAL

287	525200256	100.00		
-----	-----------	--------	--	--

2) SKILL X CONDITIONS - PRONOUN SENTENCE READING TIMES

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
RDR•SUB STRATUM					
RDR	1	13031720	7.54	13031720	3.404
RESIDUAL	22	82058976	47.47	3729953	5.137
TOTAL	23	95090688	55.01	4134378	5.690
RDR•SUB•COND STRATUM					
COND	4	12657068	7.32	3164267	4.355
RDR•COND	4	1162066	0.67	290515	0.403
RESIDUAL	88	63942720	36.99	726622	
TOTAL	96	77761840	44.99	810019	
GRAND TOTAL	119	172852512	100.00		

3) SKILL X FOREGROUNDING X SEQUENTIAL POSITION - PRONOUN SENTENCE

READING TIME

28	SOURCE OF VARIATION	DF	SS	SS%	MS	Y
30	RDR.SUB STRATUM					
	RDR	1	9434826	8.86	9434826	4.17
32	RESIDUAL	22	49774068	46.75	2262455	13.387
34	TOTAL	23	59208832	55.61	2574297	15.232
36	RDR.SUB.FOR STRATUM					
	FOR	1	5961028	4.75	5961028	11.556
38	RESIDUAL	22	9281789	8.72	421899	2.496
40	TOTAL	24	14345436	13.47	597727	3.537
42	RDR.SUB.POS STRATUM					
	POS	1	14846420	13.94	14846420	27.117
44	RESIDUAL	22	12044826	11.31	547492	3.239
46	TOTAL	24	26929156	25.29	1122348	6.630
48	RDR.SUB.FOR.POS STRATUM					
	FOR.POS	1	2184440	2.05	2184440	12.925
50	RESIDUAL	22	3718196	3.49	169009	0.519
52	TOTAL	24	5990357	5.63	249598	
54	GRAND TOTAL	95	106473760	100.00		

APPENDIX 20

ANOVA SUMMARY TABLE - EXPERIMENT 7

1) EXPERIMENTAL CONDITIONS - PRONOUN SENTENCE READING TEST

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
SUB STRATUM	17	7428456	71.72	4369685	6.036
SUB-COND STRATUM	2	4676237	4.51	2338119	3.230
RESIDUAL	34	24613764	23.76	723934	
TOTAL	36	29295000	28.28	813611	
GRAND TOTAL	53	103574656	100.00		

APPENDIX 21

ANOVA SUMMARY TABLE - EXPERIMENT 8

1) EXPERIMENTAL CONDITIONS - PRONOUN SENTENCE READING TIMES

SOURCE OF VARIATION	DF	SS	SSX	MS	VR
SUB STRATUM	17	111816352	67.94	6577432	5.427
SUB-COND STRATUM					
COND	2	11557776	7.02	5778888	4.768
RESIDUAL	34	41207832	25.04	1211595	
TOTAL	36	52765608	32.06	1465711	
GRAND TOTAL	53	164581952	100.00		

APPENDIX 22

ANOVA SUMMARY TABLE - EXPERIMENT 9

1) EXPERIMENTAL CONDITIONS X SEGMENTS

SOURCE OF VARIATION	DF	SS	SS%	PS	VR
SUB. STRATUM	23	65480112	26.16	2846561	3.673
SUB.COND STRATUM					
COND	2	21820072	8.72	10510036	6.564
RESIDUAL	46	76462288	30.54	1662224	2.144
TOTAL	48	58282352	39.26	2847549	2.641
SUB.SEG STRATUM					
SEG	1	34811952	13.91	34811952	55.202
RESIDUAL	23	14504394	5.79	831626	0.814
TOTAL	24	49316344	19.70	2154848	2.651
SUB.COND.SEG STRATUM					
COND.SEG	2	1607682	0.64	803841	1.037
RESIDUAL	46	35658048	14.24	775175	
TOTAL	48	37265728	14.89	776369	
GRAND TOTAL	143	250344512	100.00		